FAKULTÄT FÜR
ELEKTROTECHNIK,
INFORMATIK UND
MATHEMATIK

Hardwareeffiziente Echtzeit-Signalverarbeitung für synchronen QPSK-Empfang

Zur Erlangung des akademischen Grades

DOKTORINGENIEUR (Dr.-Ing.)

der Fakultät für Elektrotechnik, Informatik und Mathematik
der Universität Paderborn
vorgelegte Dissertation
von

Dipl.-Ing. Dipl.-Wirt.-Ing. Sebastian Hoffmann
aus Bielefeld

Referent:	Prof. Dr.-Ing. Reinhold Noé
Korreferent:	Prof. Dr.-Ing. Ulrich Rückert

Tag der mündlichen Prüfung: 26.06.2008

Paderborn, den 14.08.2008

Diss. EIM-E/241

Inhaltsverzeichnis

Abbildungsverzeichnis

Kapitel 1

Einleitung

1.1 Motivation

Nach turbulenten Jahren ist der Telekommunikationsmarkt wieder im Wachstum begriffen. Ein immer größerer Teil der wachsenden Weltbevölkerung nutzt immer mehr und immer aufwendigere Übertragungsdienste, und es wird nach einer Phase der Konzentration und Konsolidierung auch wieder in den Ausbau der Netzkapazität investiert. Die langsam wieder steigende Nachfrage nach höheren Übertragungskapazitäten ist allerdings gekennzeichnet durch erheblichen Kostendruck. Der Trend geht deshalb zu höheren Übertragungsraten an Stelle von neu installierten Übertragungsstrecken.

Der etablierte Langstrecken-Ethernetstandard mit einer Übertragungsrate von 10 GBit/s soll in Zukunft nicht nur durch einen 40 GBit/s-Standard abgelöst werden, sondern möglichst durch 100 GBit/s, wobei die Frage der technischen Realisierung durchaus noch ungeklärt ist.

Als sicher kann vorerst nur die allgemein steigende Nachfrage nach immer leistungsfähigeren optischen Übertragungsverfahren gelten. In Abb. 1.1 sind die technischen Details der heute absehbaren Entwicklung als Abfolge von Technologiegenerationen dargestellt: in der 1. Generation (etablierter Standard, Vergangenheit) wurden einfache Intensitätsmodulation (*OOK, on-off-keying*) und Direktempfang eingesetzt, Verfahrensdetails wie NRZ (*non-return to zero*) änderten nichts an diesem Grundprinzip. In der 2. Generation, die als marktreif gilt, wurden interferometrische Empfänger eingesetzt, die in Verbindung mit höheren Modulationsverfahren wie Quadraturphasenumtastung (*quaternary phase shift keying*, QPSK) eine optische differenzielle Dekodierung erlauben. Bei der 3.

Generation wird aller Voraussicht nach eine Leistungssteigerung gegenüber der 2. Generation dadurch erreicht, dass Überlagerungsempfänger (*Coherent detectors*, oft auch als "kohärente Empfänger" übersetzt) in Verbindung mit höherstufigen Modulationsverfahren verwendet werden. Marktreife wird für das Jahr 2010 erwartet.

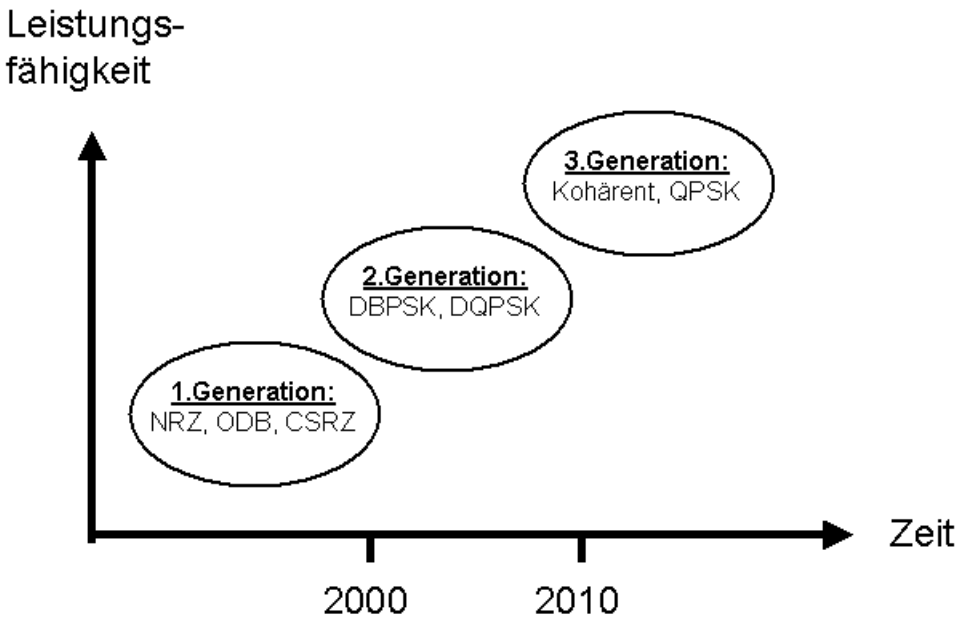

Abbildung 1.1: Technologiegenerationen in der optischen Nachrichtentechnik

Der optischen Übertragung durch Glasfasern kommt bei der erwarteten Erweiterung der Übertragungskapazitäten eine Schlüsselstellung zu, wofür es im wesentlichen zwei Gründe gibt:

1. Durch Wellenlängenmultiplex (*Wavelength Division Multiplex, WDM*) können auf einer einzigen Glasfaser zahlreiche Datenkanäle gleichzeitig übertragen werden.

2. Mit optischen Verstärkern können weite Strecken überbrückt werden, ohne dass elektronische Signalregeneration erforderlich wäre.

Bei den Kosten für den Ausbau der weltweiten Glasfasernetze dominieren die Verlegungskosten, nicht der Preis für die elektronischen und optischen Endge-

räte. Daher ist es wirtschaftlich besonders interessant, die Kapazität der bereits verlegten Glasfasern besser auszunutzen. Es gibt zwar auch noch bislang völlig ungenutzte Glasfasern, die zu Zeiten der geradezu euphorischen Wachstumserwartungen bis zum Jahr 2000 verlegt wurden, aber auch sie sollten bei ihrer Inbetriebnahme optimal ausgenutzt werden. Generell wird das Kostenargument für Ausbauentscheidungen immer wichtiger.

Ein naheliegender Ansatz zur Bandbreitenvergrößerung besteht in der Erhöhung der Symbolrate von derzeit 10 GBit/s unter Beibehaltung des Übertragungsverfahrens (digitale Intensitätsmodulation und Direktempfang). Elektronische und optische Komponenten für Übertragungsraten bis zu 40 GBit/s sind mittlerweile kommerziell verfügbar, allerdings meist zu hohen Preisen. Durch den erheblichen Kostendruck konnte sich der Markt für 40 Gbit/s-Komponenten noch nicht erholen, 10 Gbit/s ist bis heute als Standard etabliert. Auch Probleme mit den verlegten Glasfasernetzen, namentlich störende nichtlineare Effekte wie chromatische Dispersion und Polarisationsmodendispersion, verzögern den Wechsel auf eine höhere Übertragungsrate.

Eine weitere Möglichkeit zur Erhöhung der Übertragungskapazität besteht darin, das Übertragungsformat zu ändern, technisch gesehen vorrangig das Modulationsverfahren. Wenn es gelingt, die bisherige Intensitätsmodulation (OOK) mit 1-Bit-Symbolen durch ein Verfahren mit höherem Informationsgehalt pro Symbol zu ersetzen, braucht die Symbolrate[1] nicht erhöht zu werden. Das Übertragungsmedium Lichtwelle für sich genommen besitzt mit Phase und Polarisationszuständen bislang ungenutzte Freiheitsgrade, die für modernere Modulationsverfahren, wie sie in anderen Bereichen der Nachrichtentechnik bereits etabliert sind[2], genutzt werden könnten. Interessant ist dabei insbesondere die Phasenmodulation, bei der aufgrund der konstanten optischen Leistung eine geringere Empfindlichkeit gegen Nichtlinearitäten der Faser besteht und die optischen Verstärker in einem definierten Arbeitspunkt betrieben werden können.

Im Mittelpunkt dieser Arbeit steht ein solches modernes Modulationsverfahren, die Quadraturphasenumtastung mit 2 Bit/Symbol. Kombiniert mit Polarisationsmultiplex ermöglicht QPSK sogar die Vervierfachung der Bitrate gegenüber der bisherigen Intensitätsmodulation. Man kann also mit einer für die Glasfasern unkritischen Symbolrate von 10 GBaud eine Datenrate von 40 Gbit/s erreichen, was

[1] nach ihrer Einheit Baud oft auch als 'Baudrate' bezeichnet im Gegensatz zur 'Bitrate'

[2] auch das Modulationsverfahren QPSK wird seit langem eingesetzt, beispielsweise bei Telefax und Satelitenfunk

deutlich einfacher und kostengünstiger ist als eine Umrüstung auf 40 GBit/s-Direktempfang. Für die Umstellung des Modulationsverfahrens auf QPSK wird neben der senderseitigen Modulation auch eine zuverlässige Demodulation benötigt, die wiederum aus einem optischen und einem elektronischen Teil besteht.

Um Quadraturphasenumtastung verwenden zu können, ist optischer Überlagerungsempfang (kohärenter Empfang) ein dem interferometrischen Empfang überlegener Ansatz[3]. Sollen dabei kommerzielle Standardlaser nach dem DFB-Prinzip (*distributed feedback*) zum Einsatz kommen, tritt jedoch das Problem auf, dass das elektrische Empfangssignal mit einem Zwischenträger moduliert erscheint, dessen Phase durch den Frequenzunterschied der beiden Laser und ihr Phasenrauschen bestimmt wird.

Eine Verbesserung ermöglicht in diesen Fällen der Synchronempfang, bei dem die unerwünschte Zwischenträgermodulation geschätzt und bei der elektrischen Demodulation ausgeglichen wird. Die zentrale und zeitkritische Aufgabe eines synchronen Empfängers besteht darin, den Verlauf der Zwischenträgerphase so gut zu schätzen, dass eine deutliche Verbesserung gegenüber einem Asynchronempfänger erzielt wird. Zur Lösung dieser Aufgabe wurde das Konzept aus [Noe04] näher untersucht, ein diesem verwandter allgemeinerer Ansatz [Viterbi83] verbessert und schließlich ein gänzlich neuartiges und besonders hardwareeffizientes Verfahren entwickelt. Letzteres ermöglichte durch seinen geringen Platzbedarf bei der Implementation auf einem FPGA oder CMOS-Chip die weltweit erste erfolgreiche Durchführung von zwei Echtzeitübertragungsexperimenten.

1.2 Das EU-Projekt synQPSK

Die hier dargestellten Forschungsergebnisse wurden im Rahmen des seit 2004 von der Europäischen Union geförderten Projektes synQPSK erarbeitet. An der Universität Paderborn wurden die elektronischen Schaltungen entwickelt, mit denen aus dem empfangenen elektrischen Signal, das zunächst zeit- und wertkontinuierlich ist, letztendlich die gesendeten digitalen zeitdiskreten Datenströme wiedergewonnen werden müssen. Vor der eigentlichen digitalen Datenrückgewinnung muss also Taktrückgewinnung und Diskretisierung erfolgen. Eine

[3]Ein Direktempfang QPSK-modulierter Signale ist nicht möglich, weil dabei die Phaseninformation verlorengeht.

umfassende Dokumentation des Projektes ist über den Projektabschluss hinaus unter http://ontw0.upb.de/synQPSK/ verfügbar .

Im Rahmen dieser Dissertation wird vorrangig die digitale Signalverarbeitung behandelt, die behandelten Größen sind also bereits zeit- und wertediskret. Die Abtatstrate ist zugleich die Symbolrate ($1/T$, wobei T die Symboldauer ist), spezielle Probleme von Taktrückgewinnung und A/D-Umsetzung werden nicht behandelt. Die Taktrate der digitalen Signalverarbeitung ist um den Faktor M kleiner als die Abtastrate, so dass CMOS-Standardzellen aus der Bibliothek des Technologieanbieters (ST Micoelectronis) verwendet werden können. Dazu muss aus dem einfachen oder bei Polarisationsmultiplex doppelten Eingangsdatenstrom durch Demultiplex ein für M parallel arbeitende gleichartige Module verwendbares Signal erzeugt werden, wie es in [Noe05] beschrieben ist.

Die digitale Signalverarbeitung, die im folgenden Abschnitt näher beschrieben wird, lässt sich nach Umsetzung aus Matlab in die Hardwarebeschreibungssprache VHDL (*VHSIC Hardware Description Language*) sowohl in einem CMOS-Chip aus Standardzellen als auch auf einem FPGA (*Field programmable Gate Array*) synthetisieren. Dadurch wurde es möglich, die entwickelten Algorithmen nicht nur zu simulieren, sondern auch durch Versuche mit FPGAs zu testen, bevor sie als CMOS-Chip realisiert werden konnten. Allerdings war bei den FPGA-Experimenten die letztendlich angestrebte Symbolrate von 10 GBaud nicht erreichbar, was die Trägerphasenrückgewinnung erschwert [PfauCOTA].

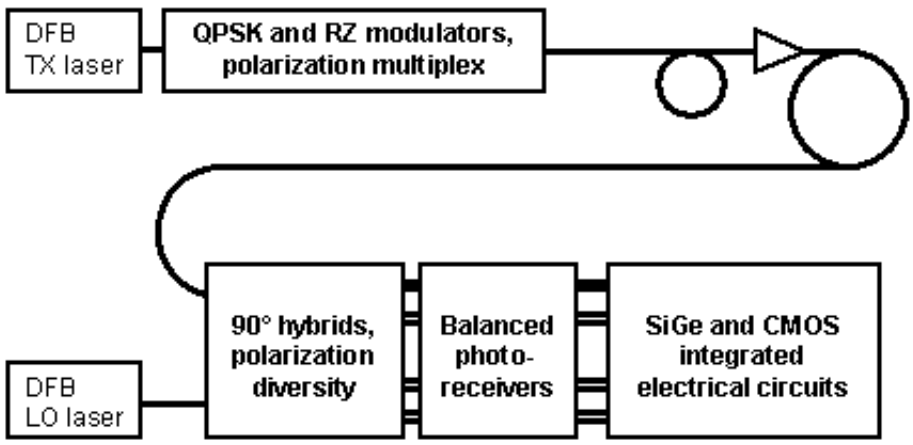

Abbildung 1.2: Schematischer Aufbau einer QPSK-Übertragungsstrecke mit zu entwickelnden Komponenten

1.3 Aufbau der Arbeit

In dieser Dissertation wird zunächst nur für eine Polarisation das Standardkonzept zur Phasen- und Datenrückgewinnung aus [Noe04] analysiert und durch Simulationen mit alternativen Konzepten verglichen. Um für den geplanten synchronen QPSK-Empfänger die Trägerphasenrückgewinnung, also das entscheidende Element, zu konstruieren, wurden zum existierenden Konzept theoretische Analysen durchgeführt, die in Kapitel 2 dokumentiert sind. Auch das mit dem Problem der Phasenrückgewinnung eng verwandte Problem der Frequenzschätzung wird in diesem Kapitel behandelt.

Beim Hauptgegenstand dieser Arbeit, der Phasenrückgewinnung, konnte gegenüber dem ursprünglich bei [Noe04, Noe05] vorgeschlagenen Konzept eine deutliche Verbesserung bei gleichzeitiger Aufwandsreduzierung erzielt werden, und zwar insbesondere mit einem verbesserten Viterbi-Phasenschätzer (Kapitel 3) und einem neuartigen und mittlerweile zum Patent angemeldeten Konzept (Kapitel 4). Die in Kapitel 3 dargestellten Verbesserungen sind theoretisch begründet und wurden durch Monte-Carlo-Simulationen überprüft. Aspekte der praktischen Umsetzbarkeit auf einem CMOS-Chip oder FPGA wurden im Simulationsmodell ebenfalls berücksichtigt.

Die Begründung des Verfahrens aus Kapitel 4 ist gegenüber Kapitel 3 eher anschaulich als theoretisch gehalten, denn es handelt sich in erster Linie um einen heuristischen Ansatz, bei dem der Aspekt der Hardwareeffizenz im Vordergrund stand. Durch die Kompaktheit der gefunden Lösung wurden eine CMOS-Realisierung und eine FPGA-Implementierung möglich. Wie bei den Verfahren aus Kapitel 3 wurden Monte-Carlo-Simulationen zum Vergleich durchgeführt. Bei der Herleitung dieses zum Patent angemeldeten und auch praktisch bewährten Konzeptes wird auch ein von der Idee her nahe verwandtes aber unbrauchbares Konzept vorgestellt.

In Kapitel 5 wird die Erweiterung des Übertragungssystems auf geregeltes Polarisationsmultiplex behandelt. Alle Konzeptvarianten zur Trägerphasenrückgewinnung, von denen nahezu 100 simuliert wurden, eignen sich grundsätzlich auch für zwei Polarisationen. Die Trägerrückgewinnung wird bei allen Konzepten besser, wenn ihr die Daten aus beiden Polarisationskanälen zugeführt werden können (nützliche Redundanz). Allerdings wird für eine gute Dekodierung auch eine saubere Trennung der beiden Polarisationen vorausgesetzt. Die in [Noe05] vorgeschlagene Regelung durch Multiplikation der Eingangsdaten mit einer zeit-

veränderlichen Kompensationsmatrix ist mit allen Konzepten kombinierbar, erhöht den Aufwand jedoch beträchtlich. Die Kompensationsmatrix muss mit Hilfe eines Korrelationsverfahrens eingestellt und laufend aktualisiert werden.

Bei der Erweiterung des Systems auf geregeltes Polarisationsmultiplex fiel die Wahl wieder auf die hardwareeffiziente Phasenrückgewinnung aus Kapitel 4, um für die neuen Komponenten (Kompensation, Korrelation und Matrizenaktualisierung) genug Platz auf FPGA bzw. CMOS-Chip zu schaffen. Die Echtzeitexperimente mit Polarisationsmultiplex [PPPPTL07] wären sonst nicht durchführbar gewesen.

Kapitel 2

Grundlagen der Phasen- und Frequenzschätzung

2.1 Überlagerungsempfang mit DFB-Lasern

Optischer Direktempfang (*Direct detection*) ist bislang das Standardverfahren in kommerziellen optischen Übertragungssystemen, weil es einfach und kostengünstig ist. Nachteilig wirkt sich dabei aus, dass als Modulationsverfahren bei Direktempfang lediglich Intensitätsmodulation (OOK, Leistungsamplitudenmodulation) möglich ist, weil der Ausgangsstrom der Empfängerdiode der Intensität proportional ist und durch Frequenz- oder Phasenmodulation nicht beeinflusst wird. Ein weiterer Nachteil des Direktempfangs gegenüber dem im Folgenden behandelten Überlagerungsempfang, nämlich die geringere Empfindlichkeit aufgrund der geringeren optischen Gesamtleistung und die daraus resultierende Begrenzung der überbrückbaren Entfernung, wurde durch leistungsfähige optische Verstärker überwunden, so dass zunächst kein kommerzielles Interesse an der Entwicklung von Überlagerungsempfängern bestand.

Überlagerungsempfang bezeichnet im Gegensatz zum optischen Direktempfang allgemein ein Verfahren, bei dem das unmodulierte Signal eines lokalen Lasers dem empfangenen Signal in einem geeigneten Koppler hinzugefügt wird, bevor es in einem differenziellen Photoempfänger gelangt. Damit das empfangene Interferenzsignal für die Dekodierung verwertbar ist, muss der lokale Laser bezüglich Moden, Polarisationen und optischer Kreisfrequenz mit dem Sender möglichst genau übereinstimmen.

Für die optische Demodulation in einem Überlagerungsempfänger benötigt man

ähnlich wie bei einem Funkempfänger eine Trägerschwingung (Laserstrahl) und den eigentlichen Demodulator, in dem das empfangene Signal mit dem Träger überlagert wird. Als optischer Demodulator eignet sich beispielsweise das 90°-Hybrid, welches von CeLight/Israel im Rahmen des Projektes synQPSK entwickelt und bei den Experimenten eingesetzt wurde. Alternativ kann auch ein preisgünstigerer 3x3-Koppler als optischer Demodulator eingesetzt werden[PHAPOpEx].

Das elektrische Empfangssignal wird erst nach dieser optischen Demodulation durch Empfängerdioden erzeugt, die in der Regel paarweise differenziell verschaltet sind. Mit zwei Empfängerdiodenpaaren und entsprechender Verschaltung der optischen Elemente (Koppler, Phasenschieber) erhält man dann ein komplexes elektrisches Empfangssignal, aus dem neben der Amplitude auch die Phase bzw. Frequenz bestimmt werden kann.

Nicht zu den Überlagerungsempfängern zählen die heute bereits marktreifen interferometrischen Empfänger für BPSK und DQPSK. Von DQPSK[1] spricht man, wenn die optisch demodulierende Trägerschwingung das um ein Symbol verzögerte Sendesignal selbst ist, das in einem Mach-Zehnder-Interferometer für eine optische differenzielle Dekodierung eingesetzt wird, weshalb man auch von einem selbsthomodynen Empfänger sprechen kann.

DQPSK-Empfang scheint zunächst besonders attraktiv, denn er kommt ohne lokalen Laser aus, womit es auch keine Intradyn-Zwischenträgerfrequenz (*intermediate frequency, IF*) geben kann, nur ein aus der Verzögerung resultierendes Phasenrauschen. Die für den Empfang verfügbare Energie des Senderlasers ist allerdings geringer als beim QPSK-Empfang mit 90°-Hybrid und lokalem Laser, weil die Strahlungsleistung des LO-Lasers nicht durch die Übertragungsstrecke abgeschwächt wird. Zu DQPSK vgl. auch [Mili05].

Vom theoretischen Standpunkt aus ist QPSK-Synchronempfang (synQPSK) besonders attraktiv, weil er aufgrund der hohen Leistung des lokalen Lasers[2] die höchste Empfindlichkeit (*Sensitivity*) aufweist. Synchroner Empfang[3] ist aus theo-

[1] wörtlich bedeutet DQPSK eigentlich nur differenziell kodiertes QPSK, was DQPSK aber nicht von dem in dieser Arbeit behandelten QPSK-System mit Überlagerungsempfang unterscheidet. Mit einem DQPSK-System ist vor allem ein spezieller Empfängertyp gemeint, in dem die differenzielle Dekodierung wie beschrieben optisch erfolgt.

[2] hoch verglichen mit der durch die Übertragungsstrecke abgeschwächten Senderlaserleistung

[3] hier zunächst in dem Sinne zu verstehen, dass die beiden Laser synchronisiert sind. Im folgenden Abschnitt bedeutet Synchronempfang die Benutzung eines *elektronischen* Phasenschätzers, ohne dass der LO im strengen Sinne (OPLL) synchronisiert wird.

retischer Sicht optimal, bedarf jedoch einer Trägerphasenrückgewinnung in Echtzeit (*real-time carrier phase recovery*).

Die Standardlösung für diese Aufgabe wäre eine Phasenregelschleife (*PLL, phase locked loop*), welche aber für den angestrebten Einsatz von DFB-Lasern (*DFB: distributed feedback*) nicht ausreicht. Nur mit ECLs *(external cavity laser)* sind bereits Anfang der 90er Jahre Versuche gelungen. Man spricht in diesem Fall auch von optischen PLLS (OPLLs), weil der gesteuerte Oszillator kein elektronischer Oszillator (VCO, NCO) sondern ein Laser ist. ECLs sind jedoch sehr aufwendige Konstruktionen und deshalb vergleichsweise teuer gegenüber DFB-Lasern. Ein kommerzieller Einsatz der Methode erschien deshalb damals unattraktiv, wie z. B. [Hooij94], S. 88 f folgert:

> Coherent IF-detection requires extremely good phase coherence between received signal and LO. When using semiconductor lasers, this can only be achieved with an (optical) PLL and very narrow linewidth (e. g. external cavity) lasers. It is above all the complexity of these small-linewidth lasers that makes receivers with coherent IF-detection unattractive at present.

Eine Besonderheit des synchronen QPSK-Empfangs bei kohärenter optischer Übertragung mit DFB-Lasern ist das Vorhandensein eines nicht vernachlässigbaren Zwischenfrequenzträgers (*IF carrier*). Der Zwischenträger resultiert einerseits aus dem Frequenzunterschied zwischen Senderlaser und lokalem Laser, andererseits aus ihren Linienbreiten und damit ihrem Phasenrauschen. Das Problem der Eliminierung des Zwischenfrequenzträgers galt damals noch, wenn nicht als unlösbar, so doch als schwierig genug um auf andere Methoden auszuweichen. So schreibt etwa [Ryu95], S. 20:

> Generally speaking, it is relatively difficult to perform synchronous detection in a coherent receiver because of the phase noise of the signal light, and for that reason asynchronous detection is preferred.

Dem Pessimismus der zitierten Lehrbücher widerspricht zwar ein bereits 1992 erfolgreich durchgeführtes und publiziertes Experiment [Noe92], aber kohärenter Empfang und Synchrondemodulation blieben für etliche Jahre ein eher exotisches Forschungsthema. Die heutige Lösung (OPLL-freie Feedforward-Lösung

mit schneller digitaler Signalverarbeitung) erschien damals noch nicht als gangbare Alternative.

In [Noe03] wurde ein analoges Konzept für einen synchronen QPSK-Empfänger vorgestellt, bei dem durch je zweifache Frequenzverdopplung und Frequenzhalbierung die Trägerfrequenz zurückgewonnen wird[4]. Trägerrückgewinnung bei einem QPSK-Signal durch zweimaliges Quadrieren anschließendes Herunterteilen der nutzsignalfreien frequenzvervierfachten Trägerschwingung findet sich z. B. auch bei [Mäusl95] S. 210, Bild 3.20. Diese als Frequenzvervielfachung und -teilung aufgefasste Trägerrückgewinnung kann prinzipiell mit einigen wenigen analogen Schaltungselementen erfolgen: Gilbertzellen [Gilb68], Addierer und Emitterfolger.

Dieser Ansatz funktioniert dann gut, wenn tatsächlich nur eine Trägerschwingung mit konstanter Frequenz wiederzugewinnen ist. Es muss aber bei DFB-Lasern aufgrund der Laserlinienbreite von einem zeitvarianten Träger ausgegangen werden, dessen momentaner Verlauf sich nicht einfach als Sinusschwingung beschreiben lässt.

In den letzten Jahren wurden immer leistungsfähigere und dabei preiswerte Komponenten für eine digitale Signalverarbeitung verfügbar. Dem zunächst gegenüber der Analoglösung höheren Schaltungsaufwand steht bei der digitalen Signalverarbeitung die zahlreichen Möglichkeiten zur späteren Erweiterung gegenüber. Als Erweiterung zu den in dieser Arbeit behandelten Elementen Phasenschätzung, Datenrückgewinnung und Polarisationsregelung könnten in Zukunft Entzerrer und Regelungsschaltungen für zahlreiche weitere Effekte wie nichtlineares Phasenrauschen, chromatische Dispersion und Polarisationsmodendispersionskompensation integriert werden.

Deswegen erfolgte der Übergang zur digitalen Lösung, die allerdings schnelle A/D-Umsetzer, Demultiplexer und einen beträchtlichen Aufwand an digitalen Schaltungselementen erfordert. In [Noe04] wurde eine digitale Umsetzung des ursprünglich analogen Konzeptes vorgestellt, die dann in [Noe05] noch auf geregeltes Polarisationsmultiplex ausgeweitet wurde.

[4] *Feedforward carrier recovery* im Titel bezeichnet den Gegensatz zu PLL-basierten Lösungen.

2.2 Mathematisches Modell des Empfangssignals

Das zu dekodierende QPSK-Signal erscheint am Empfänger mit einer Zwischenträgerfrequenz (*intermediate frequency, IF*) moduliert, die aus dem Frequenzunterschied zwischen Senderseite und Empfängerseite resultiert. Diese Zwischenträgerfrequenz ist im Idealfall gleich Null[5] muss aber in der Praxis als positiv oder negativ und außerdem zeitlich veränderlich angenommen werden. Ursache dafür sind Linienbreiten der Laser bzw. deren Phasenrauschen.

Weitere Rausch- und Verzerrungsquellen werden hier zunächst vernachlässigt. Das analoge, zeitdiskrete und komplexe Signal $z(k)$ hat dann die folgende Form:

$$z(k) = c(k)e^{j\varphi(k)} \tag{2.1}$$

wobei der Anteil $c(k) = \pm 1 \pm j$ das differenziell kodierte QPSK-Nutzsignal darstellt und während der Symboldauer als konstant gilt. Strenggenommen zeitkontinuierlich ist dagegen der Faktor $e^{j\varphi(k)}$, weil er von der zeitvarianten IF-Trägerschwingung verursacht wird. Er wird im folgenden auch als IF-Phasor bezeichnet. Da mit der Symbolrate abgetastet wird, wird auch für den IF-Phasor ein zeitdiskretes Modell verwendet, wie es z.B. bei [Hooij94], S.50f zusammenfassend beschrieben ist.

Das Spektrum der im Rahmen dieser Arbeit relevanten Lasertypen (ECL oder DFB-Laser) kann grob vereinfacht[6] durch zwei Parameter gekennzeichnet werden: Die Mittelkreisfrequenz ω und die Linienbreite Δf. Das Leistungsdichtespektrum (Lorentz-Verteilung) besitzt bei ω sein Maximum, und die Linienbreite ist die Breite des Bereichs, in dem die spektrale Leistungsdichte oberhalb des Maximalwertes minus 3 dB liegt.

Die Phase der elektrischen Feldstärke kann beschrieben werden durch einen linearen Anteil entsprechend der Mittelfrequenz ω und einen zufälligen Anteil, der umso dominanter ist, desto größer die Linienbreite Δf ist. Der lineare Anteil entspricht also einer Drehbewegung mit konstanter Geschwindigkeit, der durch Δf eine zufällige Bewegung *(random walk)* überlagert wird, die als Phasenrauschen bezeichnet wird.

[5] dieser Spezialfall wird als Homodynempfang bezeichnet, ansonsten spricht man von Intradynempfang

[6] Ein Laser mit externem Resonator (ECL) besitzt üblicherweise eine im Nahbereich verbreiterte Linie, während ein DFB-Laser mit nur einfachem Isolator eher eine im Fernbereich verbreiterte Linie besitzt.

Eine andere übliche Bezeichnung für diesen Effekt ist Frequenz-Phasenrauschen; das Phasenrauschen kann nämlich auch als Frequenzschwankung interpretiert werden. Hier soll aber ausdrücklich zwischen der deterministischen(!) Zwischenträgerfrequenz-Komponente und dem eigentlichen Phasenrauschen unterschieden werden.

Messtechnisch führt das Phasenrauschen zu einer Schwankung der momentanen Winkelgeschwindigkeit, so dass ω nur durch Mittelung über etliche Werte messbar ist (vgl. Abschnitt 2.6). Ursache des Phasenrauschens sind die einzelnen Photonen, die mit einer zufälligen Phase emittiert werden. In einem Intervall Δt ändert sich das Argument des Phasors um $\omega(\Delta t) + \Delta\varphi$, wobei $\Delta\varphi$ mittelwertfrei gaußverteilt ist mit $\sigma_{\Delta\varphi} = \sqrt{2\pi(\Delta f)(\Delta t)}$. Die Wahrscheinlichkeitsdichtefunktion p für $\Delta\varphi$ mit dem Parameter Δt lautet (vgl. [Hooij94], S.51, Formel (2-29)):

$$p(\Delta\varphi, \Delta t) = \frac{\exp\left(-\frac{(\Delta\varphi)^2}{4\pi(\Delta f)(\Delta t)}\right)}{\sqrt{4\pi^2(\Delta f)(\Delta t)}} = \frac{e^{-\frac{(\Delta\varphi)^2}{2\sigma^2_{\Delta\varphi}}}}{\sqrt{2\pi}\sigma_{\Delta\varphi}} \tag{2.2}$$

Die etwas ungenaue Gleichsetzung von Parameter und freier Variable bei Wahrscheinlichkeitsdichtefunktionen wird hier originalgetreu beibehalten. Für das zeitdiskrete Simulationsmodell mit einem Wert pro Symbol ist $\Delta t = T$ zu setzen, also der Kehrwert der Symbolrate; für ein Modell mit Überabtastung wäre für Δt ein entsprechender Bruchteil von T zu setzen.

Auch für den IF-Phasor, der ja aus der Überlagerung zweier Laserspektren und der inhärenten Tiefpassfilterung durch die elektronischen Elemente resultiert, gilt die vorstehende Beschreibung. Beim Überlagerungsempfang wirkt das 90°-Hybrid als optischer Demodulator, die zeitabhängigen Terme $e^{j\varphi}$ werden miteinander multipliziert.

So entsteht im wesentlichen eine Trägerschwingung bei der Differenzfrequenz der beiden Laser, die als Zwischenträgerfrequenz bezeichnet wird. Da sowohl Senderlaser (Index S) als auch der lokale Laser des Überlagerungsempfängers (Index LO) durch die beiden Parameter gekennzeichnet sind, lautet der IF-Trägerphasor:

$$e^{j\varphi} = e^{j((\omega_S - \omega_{LO})t + \Delta\varphi_S - \Delta\varphi_{LO})} = e^{j(\omega_{IF}t + \Delta\varphi_{IF})} \tag{2.3}$$

Auch φ, die Zwischenfrequenzphase, ist also aus einem konstanten Frequenzanteil ω_{IF} und einem *Randomwalk*-Anteil zusammengetzt. Dementsprechend werden im Simulationsprogramm keine separaten Modelle für die beiden Laser S

und LO verwendet, sondern es wird direkt eine Rekursionsformel für den IF-Phasor mit den Parametern $\Delta\varphi_o = \omega_{IF} \cdot T$ (*dphioffset*, entsprechend der IF relativ zur Symbolrate) und dem Quotienten aus Linienbreite und Symbolrate (*LinewidthtimesT*). Aus letzterem wird der Parameter $\sigma_{\Delta\varphi}$ (*sigmadphi*) berechnet[7].

Die für die Beschreibung wesentliche Rauschquelle ist die Glasfaser-Übertragungsstrecke, in der thermisches Rauschen mit einer Leistung entsprechend der Länge hinzugefügt wird. Gaußverteiltes additiv überlagertes Rauschen (AWGN, *additive white Gaussian noise*) bei Real- und Imaginärteil wird in der Simulation anhand eines vorgegebenen Signal-Rauschverhältnises (*signal to noise ratio*, SNR) erzeugt und dann dem Signal als komplexer Rauschphasor $r(k)$ hinzugefügt. Somit ergibt sich für die empfangenen Symbole $z(k)$ die einfache zeitdiskrete wertkontinuierliche Beschreibung:

$$z(k) = c(k) \cdot e^{j\varphi(k)_k} + r(k) \tag{2.4}$$

Das ganzzahlige Argument k bezeichnet zunächst nur die zeitliche Reihenfolge der Abtastwerte. Durch Demultiplexer und Speicherelemente sind zu einem Abtastwert $z(k)$ sowohl einige Vorgänger als auch einige Nachfolger verfügbar. Bei der Bezeichnung der aus den $z(k)$ berechneten Größen wird aus Gründen der Übersichtlichkeit der Index beibehalten, z. B. bei $\psi(k) = \text{arc}(z(k))\text{mod}2\pi$.

Dass $\psi(k)$ erst berechnet oder aus einer Tabelle ausgelesen werden muss und deshalb erst später zur Verfügung steht als $z(k)$, spielt in dieser Notation keine Rolle, weil Probleme mit unterschiedlichen Auslese- und Berechnungszeiten in der Praxis durch Einfügung von Verzögerungselementen leicht zu kompensieren sind. Diese Einführung von Verzögerungselementen bildet die Grundlage für Vorwärtskopplung (*feed forward*), bei denen auch ein der Beschreibung nach nicht kausales Verhalten, nämlich der Zugriff auf 'zukünftige' Werte, möglich wird.

Bislang wurden Rauschquellen und Verzerrungseffekte innerhalb der Übertragungsstrecke noch nicht behandelt. Nicht ideales Verhalten zeigen unter anderem Modulator, 90°-Hybrid, Empfängerdioden und ADCs. Im Rahmen dieser Arbeit werden nur einige besondere Effekte der AD-Umsetzung ausführlich behandelt, obgleich im modular aufgebauten Simulationsmodell noch weitere nichtlineare Effekte berücksichtigt sind.

[7] Der entsprechende Teil des Matlab-Simulationsprogrammes lautet:
sigmadphi=sqrt(2*pi*LinewidthtimesT); IFphasor=ones(1,itmax);
dphi=real(exp(j.*2.*pi.*rand(1,itmax)).*sigmadphi.*sqrt(-2.*log(rand(1,itmax))))+dphioffset;
for it=2:itmax, IFphasor(it)=IFphasor(it-1)*exp(j*dphi(it)); end;

2.3 Differenzielle Kodierung

Ein grundsätzliches Problem bei digitaler Trägerphasenmodulation ist, dass die gesendeten Symbole (anschaulich dargestellt als Punkte im Konstellationsdiagramm) nach der Übertragung um einen unbekannten und i. A. zeitvarianten Winkel (Phasenoffset) verdreht erscheinen, so dass aus der absoluten Lage der empfangenen Symbole im Konstellationsdiagramm nicht eindeutig das gesendete Symbol wiedergewonnen werden kann. Das Problem lässt sich lösen, in dem man die Information nicht mit der absoluten Lage eines QPSK-Symbols kodiert, sondern mit seiner relativen Lage zum Vorgänger. Dieses Verfahren wird als differenzielle Kodierung (*differential encoding*) bezeichnet.

Es sei $n_o(k)$ eine Folge natürlicher Zahlen, die Elemente der Menge $\{0, 1, 2, 3\}$ sind. Im Hinblick auf die Identität $j^n = j^{n+4}$ für $n \in \mathbb{N}$ werden derartige Zahlen im folgenden auch als Quadrantenzahlen bezeichnet. Bekanntlich können auch endliche Teilmengen der natürlichen Zahlen mit entsprechend definierten Operationen $\oplus$ und $\odot$ Zahlenkörper bilden. Die Operationen $\oplus$ und $\ominus$ bezeichnen im folgenden Addition bzw. Subtraktion modulo 4, abgekürzt mod4.

Die Zahlen $n_o(k)$ besitzen jeweils einen Informationsgehalt von 2 Bit und lassen sich durch zweistellige Binärzahlen darstellen. Vor der Umwandlung in komplexe Sendesymbole $c(k)$ werden die Quadrantenzahlen $n_o(k)$ zunächst durch Addition des bereits gesendeten Vorgängerwertes $n_t(k-1)$ differenziell vorkodiert:

$$n_t(k) = n_o(k) \oplus n_t(k-1) \tag{2.5}$$

Als Startwert der gesendeten (*t=transmitted*) Quadrantenzahlen wird $n_t(0) = 0$ gesetzt. Die Umwandlung der differenziell vorkodierten Quadrantenzahl $n_t(k)$ in das in Gl. (2.1) und Gl. (2.4) auftretende komplexe Sendesymbol $c(k)$ erfolgt nach der Gleichung

$$c(k) = \sqrt{2} e^{j\frac{\pi}{4}} j^{n_t(k)} \tag{2.6}$$

Real- und Imaginärteil von $c(k)$ werden im QPSK-Modulator getrennt voneinander verwendet (Inphase- und Quadraturanteil). Da sich im Konstellationsdiagramm benachbarte Punkte nur jeweils durch Real- oder Imaginärteil voneinander unterscheiden, spricht man auch von einer Gray-Kodierung der Quadrantenzahlen $n_t(k)$ analog zur binären Gray-Kodierung von $(0, 1, 2, 3) = (00, 01, 11, 10)$,

die bezüglich der Hammingdistanz dieselbe Eigenschaft besitzt und hier *nicht* gemeint ist.

Es wird zunächst ein Empfänger und Entscheider angenommen, der in der Lage ist, eine Folge von empfangenen Quadrantenzahlen $n_r(k)$ (*r=received*) zu liefern, die die Eigenschaft $n_r(k) = n_t(k) \oplus n_x$ besitzen, wobei n_x eine unbekannte aber zeitlich konstante Quadrantenzahl ist.

Die differenzielle Dekodierung der Folge $n_r(k)$ wird ähnlich realisiert wie die Kodierung:

$$\hat{n}_o(k) = n_r(k) \ominus n_r(k-1) \tag{2.7}$$

Da $\hat{n}_o(k) = n_t(k) \oplus n_x \ominus (n_t(k-1) \oplus n_x) = n_o(k) \oplus n_t(k-1) \ominus n_t(k-1) = n_o(k)$ ist, erhält man bei fehlerfreier Übertragung die Originalquadrantenzahlen unabhängig von n_x zurück[8].

Ein Nachteil der differenziellen Kodierung ist, dass die Verfälschung eines übertragenen Symbols bei der anschließenden Dekodierung zu Fehlern bei zwei aufeinanderfolgenden dekodierten Symbolen führt, so dass sich bei selten auftretenden Fehlern die BER (*Bit error ratio*, Bitfehlerverhältnis) annähernd verdoppelt. Dies ist eine wesentliche Ursache für die Abweichung zwischen theoretischer und idealer BER-Kurve bei den Simulationen in [HofCOTA].

In der Dekodierergleichung (2.7) wurde vorausgesetzt, dass eine Einrichtung für Demodulation und Entscheidung, die die empfangenen Symbole liefert, dem Dekodierer vorgeschaltet ist. Es ist aber auch möglich, die Reihenfolge der Rechenoperationen zu ändern. Ein derartiger Empfänger wird als asynchroner oder differenzieller Empfänger bezeichnet, weil die Umkehrung der differenziellen Kodierung vor der Entscheidung erfolgt, ohne dass synchron demoduliert werden muss. Das verringert den Hardwareaufwand beträchtlich.

Es seien $n_t(k)$ differenziell kodierte QPSK-Quadrantenzahlen, die nun als komplexe Zahlen $z(k)$ mit unbekannten Beträgen $|z(k)| > 0$ und Lagewinkeln $\vartheta(k)$ dargestellt werden: $z(k) = |z(k)|\, j^{n_t(k)} e^{j\vartheta(k)}$. Dies entspricht verallgemeinert der Situation im Empfänger vor dem Entscheider mit idealer Übertragung nach Gl. (2.1).

[8] auf die Unterscheidung von senderseitigen $n_o(k)$ und wiederhergestellten $\hat{n}_o(k)$ wird im Folgenden in Anlehnung an [Noe04, Noe05] verzichtet. Der Index o steht wahlweise für *original* oder *output*

Die Lagewinkel zweier aufeinanderfolgender Symbole seien annähernd gleich, also $\vartheta(k) \approx \vartheta(k-1)$. Diese Annahme besagt, dass die Zwischenträgerfrequenz annähernd verschwindet und auch das Phasenrauschen gering ist. Man kann folgendes Produkt bilden, formal normieren und für den Entscheider weiterverwenden:

$$\frac{z(k)z^*(k-1)}{|z(k)|\,|z(k-1)|} = j^{n_t(k)-n_t(k-1)} \cdot e^{j(\vartheta(k)-\vartheta(k-1))} \hat{\approx} j^{n_t(k)-n_t(k-1)} = j^{n_o(k)} \tag{2.8}$$

Das Symbol $\hat{\approx}$ soll verdeutlichen, dass der Entscheider durch Rundung den Faktor $e^{j(\vartheta(k)-\vartheta(k-1))}$ eliminiert, also das normierte Produkt nur noch einem der vier Punkte $\{1, j, -1, -j\}$ zuordnet; es entsteht ein Konstellationsdiagramm, bei dem die Symbole auf den Achsen liegen. Die Beträge $|z(k)|$ und $|z(k-1)|$ spielen für die Entscheidung keine Rolle, weshalb die formale Normierung aus Gl. (2.8) nicht in Hardware umgesetzt zu werden braucht.

Die differenzielle Kodierung wird also durch die konjugiert-komplexe Multiplikation ebenso rückgängig gemacht wie durch die Differenzbildung in Gl. (2.7). Die Lagewinkel tauchen nur noch als (geringe) Differenz auf, die beim Entscheidungsvorgang zusammen mit dem Rauschen eliminiert wird. Dieses Prinzip wird optisch beim DQPSK-Empfang verwendet, wobei die konjugiert-komplexe Multiplikation im Interferometer realisiert wird.

Zunächst erscheint die komplexe Multiplikation in Gl. (2.8) komplizierter als eine Differenzbildung natürlicher Zahlen in Gl. (2.7), aber man kann betragsfrei in Polarkoordinaten rechnen, wodurch sich die komplexe Multiplikation zu einer reellen Differenzbildung vereinfacht. Dazu wird zunächst die Winkelgröße $\psi(k)$ definiert durch

$$\psi(k) := \text{arc}(z(k)) \text{mod} 2\pi \tag{2.9}$$

Der Zusatz mod2π ist in diesem Zusammenhang nicht unbedingt nötig, man kann auch den Hauptwert des Argumentes verwenden. Mit (2.9) erhält man aber stets einen positiven Winkel, der sich zerlegen lässt in den bereits erwähnten Lagewinkel $\vartheta(k)$ des empfangenen Symbols innerhalb des jeweiligen Quadranten und eine Quadrantenzahl $q(k) \in \{0, 1, 2, 3\}$ zur Bezeichnung desselben:

$$\vartheta(k) := \text{arc}(z(k)) \text{mod} \frac{\pi}{2} \tag{2.10}$$

$$q(k) := \left\lfloor \frac{2}{\pi} \text{arc}(z(k)) \text{mod} 2\pi \right\rfloor \tag{2.11}$$

$$\psi(k) = q(k)\frac{\pi}{2} + \vartheta(k) \tag{2.12}$$

Für das Argument des beim asynchronen Empfänger verwendeten komplexen Produktes nach Gl. (2.8) gilt:

$$\text{arc}\ (z(k)z^*(k-1)) \text{mod} 2\pi = (\psi(k) - \psi(k-1)) \,\text{mod} 2\pi \tag{2.13}$$

Die Entscheidung für ein $n_o(k)$ auf der Grundlage dieser Winkeldifferenz ist sehr einfach:

$$n_o(k) = \left\lfloor \frac{2}{\pi}(\psi(k) - \psi(k-1)) + \frac{1}{2} \right\rfloor \text{mod} 4 \tag{2.14}$$

Ein auf dieser Grundlage arbeitender differenzieller elektronischer QPSK-Empfänger kommt also ohne Trägerphasenrückgewinnung aus und ist damit auch kein Synchronempfänger, deshalb wird er in dieser Arbeit als asynchroner Empfänger[9] bezeichnet. Man kann ihn als elektronische Alternative zum DQPSK-Empfängers ansehen, bei dem die Operation (2.8) optisch realisiert wird. Bei Leistungsbewertungen von synchronen Empfängern stellt er ein praktisch realisierbares Vergleichssystem dar, weshalb er z. B. auch bei [Leven06] zum Vergleich herangezogen wird.

2.4 Synchronempfang

Das Prinzip der differenziellen Dekodierung ist auch bei einem Synchronempfänger anwendbar. Der im vorigen Abschnitt vorgestellte asynchrone Empfänger liefert dann schlechte Ergebnisse, wenn die Annahme $\vartheta(k) \approx \vartheta(k-1)$in (2.8) nicht mehr gut genug erfüllt ist, denn dann kommt es vermehrt zu Fehlern des Entscheiders durch die schnellen Lageveränderungen des Konstellationsdiagrammes. Bei der Synchrondemodulation wird daher in geeigneter Form

[9]nach der Nomenklatur bei [Kamm96], S.440 wäre es ein inkohärenter differenzieller Demodulator. Der Begriff Kohärenz wird leider nicht einheitlich verwendet; in jüngeren englischsprachigen Veröffentlichungen bedeutet *coherent detection* meist die in dieser Arbeit behandelte Kombination von optischem Überlagerungsempfang und elektronischer Synchrondemodulation mit einem zurückgewonnenen Träger, wie z. B. bei [IpKahn05] klargestellt wird.

die Trägerphase $\hat{\varphi}$ wiedergewonnen und das Konstellationsdiagramm mit einer entsprechenden Drehung durch Multiplikation mit $e^{-j\hat{\varphi}}$ so verändert, dass der Lagewinkel nicht mehr so stark variiert.

Wie beim Asynchronempfänger kann allein aus der Winkeldifferenz aufeinanderfolgender Werte durch Rundung entschieden werden, welches Symbol in differenzieller Kodierung gesendet wurde. Da der Betrag von $ze^{-j\hat{\varphi}}$ bei der Entscheidung keine Rolle spielt, kann die komplexe Multiplikation entfallen, man verwendet statt arc$(ze^{-j\hat{\varphi}})$ die Winkeldifferenz $\psi - \hat{\varphi}$ [Noe04].

Der geschätzte Phasenwinkel wird unabhängig vom verwendeten Schätzverfahren in dieser Arbeit so definiert, dass er auch den ursprünglichen Lagewinkel des Symbols gemäß Gl. (2.6) umfasst und zu einem in den ersten Quadranten gedrehten Symbol passt:

$$\hat{\varphi}(k) \approx \left(\varphi(k) + \frac{\pi}{4}\right) \bmod \frac{\pi}{2} \tag{2.15}$$

Wie schon beim Asynchronempfänger empfiehlt es sich auch bei der Synchrondemodulation, komplexe Multiplikationen durch Summen bzw. Differenzen der Argumente zu ersetzen, zumal der Betrag des Produktes für den nachfolgenden Entscheider ohnehin irrelevant ist. Demzufolge werden beim Synchronempfänger die empfangenen Quadrantenzahlen wie folgt gebildet:

$$n_r(k) = \left\lfloor \frac{2}{\pi}(\psi(k) - \hat{\varphi}(k)) + \frac{1}{2} \right\rfloor \bmod 4 \tag{2.16}$$

Ähnlich wie beim Entscheider des Asynchronempfängers nach Gl. (2.14) wird also eine Quadrantenzahl gebildet, die einem von vier Punkten des auf die Achsen gedrehten Konstellationsdiagramms entspricht. Da die Zuordnung (Entscheidung) vor der differenziellen Dekodierung getrennt für $n_r(k)$ und $n_r(k-1)$ erfolgt, kann beim Synchronempfänger mehr Rauschen eliminiert werden als beim Asynchronempfänger.

Das zentrale Element des Synchronempfängers ist der Phasenschätzer. Beim Originalkonzept [Noe04] wird er durch eine komplexe Mittelwertbildung zweifach quadrierter Eingangswerte realisiert, was mit modifizierter Notation in Gl. (2.17) wiedergegeben ist. Die Klammersetzung verdeutlicht die drei Hauptelemente dieser Berechnung: Potenzierung, Mittelung und Argumentbestimmung.

$$\hat{\varphi}(k) = \frac{1}{4}\left[\text{arc}\left(\frac{1}{2N+1}\sum_{n=k-N}^{k+N}\left\{z^4(n)\right\}\right)\text{mod}2\pi\right] \tag{2.17}$$

Die Realisierung des Phasenschätzers nach Gl. (2.17) wird im Rahmen dieser Arbeit als Originalkonzept bezeichnet. Auch nach einigen Umformungen zur Hardwareersparnis (Weglassen des Faktors $\frac{1}{2N+1}$ und Integration des Faktors $\frac{1}{4}$ in die arc-Tabelle) ist der Aufwand für die Realisierung des Originalkonzeptes beträchtlich: es muss komplex potenziert werden (innere geschweifte Klammern) und durch komponentenweises Aufsummieren (runde Klammern) ein Wert $y(k)$ erzeugt werden, der dann mit Hilfe einer Tabelle $\hat{\varphi}(y) = \frac{1}{4}(\text{arc}(y(k))\text{mod}2\pi)$liefert.

Der geschätzte Phasenwinkel $\hat{\varphi}(k)$ nach Gl. (2.15) und (2.17) liegt wie der Lagewinkel $\vartheta(k)$ stets im Intervall $[0, \frac{\pi}{2}[$. Andere Definitionen sind möglich und gebräuchlich (vgl. [Noe04, Noe05]), hier werden aus Gründen der übersichtlichen Hardwareumsetzung negative Winkel vermieden.

Eine wichtige Kenngröße des Phasenschätzers ist die Konstante N, die angibt, wieviele Vorgänger und Nachfolger bei der gleitenden Durchschnittsbildung nach Gl. (2.17) einbezogen werden. Umso größer N gewählt wird, desto stärker ist die Glättung und Rauschunterdrückung; die maximal mögliche Änderung zwischen zwei direkt aufeinanderfolgenden Werten $\hat{\varphi}(k)$, $\hat{\varphi}(k+1)$ wird dadurch aber auch verringert, so dass der Phasenschätzer schnellen Änderungen der physikalischen Phase $\varphi(t)$ nicht mehr folgen kann.

Der Asynchronempfänger kann als primitiver Spezialfall des Synchronempfängers betrachtet werde, bei dem $N = 0$ ist, so dass gemäß Gl. (2.17) der gemessene Lagewinkel selbst als Schätzwert für den 'richtigen' Lagewinkel verwendet wird, also die Zuweisung $\hat{\varphi}(k) := \vartheta(k)$ gilt. Ein Synchronempfänger i. e. S. verwendet dagegen weitere Informationen, um für die Phase einen besseren Schätzwert als der Asynchronempfänger zu generieren.

2.5 Sprungzahlen und Unwrapping

Die geschätzte Phase stammt auf Grund des Vorfaktors $1/4$ und der modulo-Operation stets aus dem Intervall $[0, \frac{\pi}{2}[$ und wird deshalb im folgenden auch als Einquadrantenphase bezeichnet. Eine solche Einschränkung gegenüber der physikalischen Phase, die beliebige reelle Werte annehmen kann, ist notwendig, um bei der Umkehrung der Operation $z^4(k)$ Eindeutigkeit zu erzielen.

Physikalisch korrekter wäre es, jeweils alle vier möglichen Lösungen von $\sqrt[4]{y(k)}$ zu betrachten und daraus den wahrscheinlichsten Verlauf des IF-Phasors durch alle vier Quadranten[10] zu rekonstruieren. Das ist in Echtzeit schwierig, aber nicht unmöglich. Eine Operation zur Umwandlung einer Einquadrantenhase (anschaulich als 'aufgewickelt', engl. *wrapped* bezeichnet) in eine der physikalischen Realität nähere Vierquadrantenphase wird auch als *Unwrapping* bezeichnet.

Bei [Noe04] wird die Einquadrantenphase nach Gl. (2.17) parallel zur Verwendung bei der Gewinnung der $n_r(k)$ auch dazu verwendet, Sprungzahlen (*quadrant jump numbers*) $n_j(k)$ zu generieren. Diese werden bei der nachfolgenden differenziellen Dekodierung, einer einfach modulo 4-Operation, dergestalt mit einbezogen, dass die physikalisch unwahrscheinlichen Sprünge der geschätzten Einquadrantenphase nicht zu falschen Dekodierergebnissen führen.

Das Konzept der sprungzahlgestützten Dekodierung wird in dieser Arbeit für alle Phasenschätzervarianten grundsätzlich beibehalten. Die Definition der Sprungzahl lautet:

$$n_j(k) := -\text{sign}(\varphi(k) - \hat{\varphi}(k-1)) \left\lfloor \frac{\pi}{4} |\hat{\varphi}(k) - \hat{\varphi}(k-1)| \right\rfloor \tag{2.18}$$

Diese Definition liefert ein $n_j(k) \in \{-1, 0, 1\}$, $n_j(k)$ ist also keine Quadrantenzahl im bisher gebrauchten Sinne, weil sie vorzeichenbehaftet ist. Diese Definition ist für die physikalische Interpretation (Vorzeichen als Richtung des Quadrantensprunges) hilfreich. Für die Hardwareumsetzung ist die Unterscheidung von -1 und $3 = -1 \text{mod} 4$ bedeutungslos, denn beide werden für die weitere Verarbeitung binär durch 11 dargestellt.

Es ist bemerkenswert, dass eine Sprungzahl mit $|n_j(k)| \geq 2$ definitionsgemäß nicht auftreten kann. Dies hängt mit der durch den internen Aliaseffekt (vgl. 3.2) begrenzten maximal detektierbaren Frequenz zusammen. Die durch die Einbeziehung der Sprungzahl modifizierte Dekodierergleichung lautet:

$$n_o(k) = n_r(k) \ominus n_r(k-1) \oplus n_j(k) \tag{2.19}$$

Wird das demgegenüber aufwendige und zeitkritische *Unwrapping* der geschätzten Phase gefordert, so werden zunächst ebenfalls aus der Einquadrantenphase

[10] Auch wenn die physikalische Phase jeden reellen Wert annehmen kann, so genügt zu ihrer Schätzung die Beschränkung auf ein 2π breites Intervall, weil $e^{j\varphi}$ periodisch ist.

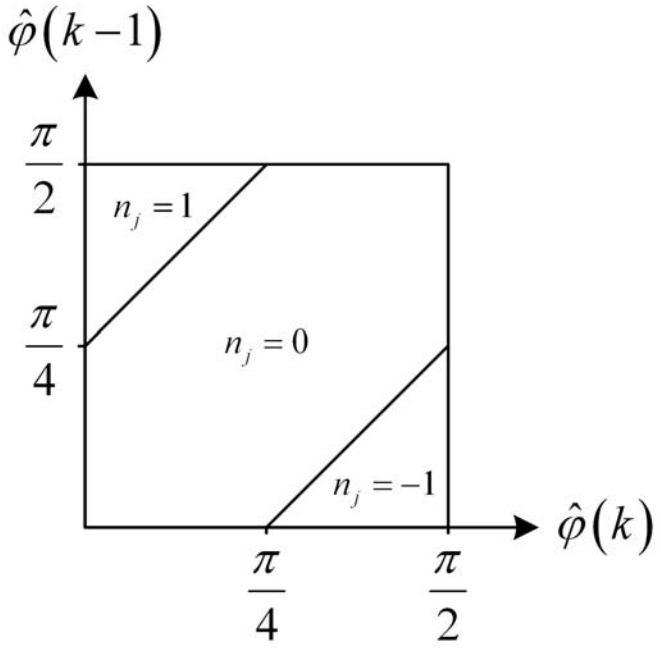

Abbildung 2.1: Sprungzahlbereiche

$\hat{\varphi}(k)$ Sprungzahlen $n_j(k)$ gebildet. Ein einzelner Wert $\hat{\gamma}(k)$ der gesuchten Vierquadrantenphase[11] lässt sich wie folgt rekursiv aus seinem jeweiligen Vorgänger definieren:

$$\hat{\gamma}(k) = \hat{\gamma}(k-1) - \hat{\varphi}(k-1) + \hat{\varphi}(k) + \frac{\pi}{2} n_j(k) \tag{2.20}$$

Die letzten drei Summanden drücken dabei die Gesamtphasenänderung gegenüber dem Vorgängerwert aus. Vereinfacht durch Zusammenfassung von drei Summanden zu einer Bezugsquadrantenzahl $n_b(k)$ lautet die Gleichung

$$\hat{\gamma}(k) = \hat{\varphi}(k) + \frac{\pi}{2} n_b(k) \tag{2.21}$$

Die Bezugsquadrantenzahl $n_b(k)$ lässt sich ebenfalls rekursiv aus ihrem Vorgänger ableiten, indem man sie modulo 4 mit der aktuellen Sprungzahl addiert:

$$n_b(k) = n_b(k-1) \oplus n_j(k) \tag{2.22}$$

Ein Zugriff auf den unmittelbaren Vorgänger ist auch bei derart einfachen Rechenoperationen noch zeitkritisch. Um die Bezugsquadrantenzahl anders zu ge-

[11] Alle Gleichungen mit $\hat{\gamma}$ sind mod2π zu verstehen. Anders als beim Hauptwert erhält man also keine negativen Winkel.

winnen, wird zunächst eine über l Werte aggregierte Sprungzahl $m_j(k,l)$ (*m: master*) wie folgt definiert:

$$m_j(k,l) := \sum_{i=k-l}^{k} n_j(i) \tag{2.23}$$

Für eine zeitunkritische Bildung von Bezugsquadrantenzahlen, die die Einquadrantenphase zur Vierquadrantenphase ergänzen, genügt es, die aggregierte Sprungzahl modulo 4 zu bilden. Dies gilt auch für die Zwischensummen bei einem zur Sprungzahlaggregation konstruierten Addiererbaum. Bei der parallelen Berechnung von M aggregierten Sprungzahlen tauchen zahlreiche gemeinsam nutzbare Zwischensummen auf.

Die Definition der Bezugsquadrantenzahl nach (2.22) wird mit Hilfe der aggregierten Sprungzahl wie folgt verallgemeinert:

$$n_b(k) = n_b(k-l) \oplus m_j(k,l) \tag{2.24}$$

Da man l beliebig groß wählen kann und die Bildung der $m_j(k,l)$ wenig Aufwand erfordert, ist die Bildung der Bezugsquadrantenzahlen nun nicht mehr zeitkritisch.

Für die Zusammenfassung der Bezugsquadrantenzahl mit der Einquadrantenphase zu einer Vierquadrantenphase muss die Einquadrantenphase $\hat{\varphi}(k)$ allerdings verzögert werden, so dass die Synchrondemodulation mit einer Vierquadrantenphase die geringe Vereinfachung der Dekodiergleichung (2.19) zu (2.7) mit einem erhebelichen Mehraufwand erkaufen würde. Deshalb wurde normalerweise bei Simulationen und Experimenten die Einquadrantenphase beibehalten.

2.6 Frequenzschätzung und Frequenzregelung

Das Prinzip der Vorwärtskopplung und Phasenschätzung[Noe03] ersetzt die problematische OPLL-Lösung nach [Derr], aber sie setzt voraus, dass tatsächlich Intradynempfang mit einem geringen Frequenzunterschied der beiden Laser vorliegt, so dass das Phasenrauschen dominiert. Ist der Frequenzunterschied zwischen Sender- und LO-Laser zu groß, so vermag die Trägerphasenrückgewinnung der physikalischen Phase nicht mehr zu folgen. Als kritisch wurden für die

in dieser Arbeit näher beschriebenen Phasenschätzer bei einer Symbolrate von 10 GBaud Werte von ca. 100 MHz durch Simulation festgestellt.

Daher wurde für die Experimente entschieden, für den LO-Laser eine grobe Frequenzregelung vorzusehen. Grob bedeutet dabei, dass weder Geschwindigkeit noch Genauigkeit besonders kritisch sind, weil diese Regelung der für die Demodulation entscheidenden Trägerphasenrückgewinnung (die man auch als Feinregelung bezeichnen könnte, weil sie die für die Demodulation kritischen rauschbedingten Schwankungen des Lagewinkels ϑ ausregelt) lediglich überlagert ist.

Diese Vorgehensweise setzt allerdings voraus, dass man in der digitalen Signalverarbeitung einen Schätzwert $\hat{f} \approx \frac{\omega_{IF}}{2\pi}$ für die Intradyn-Zwischenträgerfrequenz gewinnen kann. Dazu wird in diesem Abschnitt ein Ansatz vorgestellt, der wegen seines minimalen Hardware-Aufwandes implementiert und erfolgreich im Experiment eingesetzt wurde. Auch auf dem im November 2007 als Layout fertiggestellten "CMOS-Chip B", auf dem die Polarisationsregelung integriert ist, wurde dieser Frequenzschätzer integriert, vgl. [Wörde07] S.80f.

Wird die aggregierte Sprungzahl nach (2.23) nicht modulo 4 berechnet, sondern zu einer ganzen Zahl aufsummiert, so hat deren Betrag die Bedeutung der netto im betrachteten Zeitraum von einer Vierquadrantenphase durchlaufenen Quadranten, und das Vorzeichen gibt den Drehsinn an (eine positive Summe bedeutet mathematischen Drehsinn). Da das Phasenrauschen und die anderen Rauscheinflüsse als mittelwertfrei angenommen werden können, muss diese Bewegung (lineare Trendkomponente der Phase) Ausdruck der zu schätzenden Intradyn-Zwischenfrequenz sein. Man kann die aggregierte Sprungzahl direkt in eine geschätzte Frequenz umrechnen:

$$\hat{f}(k,l) := \frac{1}{8lT} m_j(k,l) \tag{2.25}$$

Dieser Schätzwert ist ein vorzeichenbehafteter Momentanwert. Die vereinfachende Aussage $\omega_{IF} \propto \sum n_j$ aus [Wörde07], S.80 ist also zu relativieren. Insbesondere gilt es auch die mit dem im folgenden Kapitel noch näher beschriebenen 'internen Aliaseffekt' zusammenhängende maximale detektierbare Frequenz zu beachten.

Zur Bestimmung dieser maximal detektierbaren Frequenz und Herleitung des Vorfaktors $\frac{1}{8lT}$ wird zunächst eine aus nur zwei Werten aggregierte Sprungzahl betrachtet. Man kann zeigen[12], dass zwei zeitlich benachbarte Sprungzahlen niemals den gleichen von Null verschiedenen Wert annehmen, so dass gilt:

[12] durch Rückgriff auf die Definitionen und eine Fallunterscheidung

$$m_j(k,2) = n_j(k) \oplus n_j(k-1) \in \{-1, 0, 1\} \tag{2.26}$$

Falls $|m_j(k,2)| = 1$ ist, so bedeutet dies einen Quadrantenwechsel in der Zeit $lT = 2T$. Für diesen Quadrantenwechsel hat sich $\hat{\varphi}$ einmal um mindestens $\frac{\pi}{4}$ geändert, sonst wäre keine Sprungzahl erzeugt worden. Die zweite Änderung in unbekannter Richtung hat diesen Wert vergrößert oder reduziert, aber nicht ganz aufgehoben, sonst wäre die aggregierte Sprungzahl Null. Erwartungswert für die mittlere mit dem Quadrantenwechsel verbundene Phasenänderung ist $\frac{\pi}{4}$, was einer momentanen Kreisfrequenz von $\frac{\pi}{4lT}$ entspricht. Division dieser Kreisfrequenz mit 2π liefert den Vorfaktor in Gl. (2.25). Für $l = 2$ kann die Gesamt-Phasenänderung nicht größer werden als $\frac{\pi}{2}$, was einer maximal detektierbaren Frequenz von $f_{max} = \frac{1}{8T}$ entspricht.

Diese gegenüber dem Abtasttheorem um den Faktor 4 kleinere Maximalfrequenz ist charakteristisch für die Phasenschätzung beim Modulationsverfahren QPSK, bei dem eigentlich vier gleichrangige Lösungen für jeden Wert $\hat{\varphi}(k)$ existieren und der jeweils richtige nach dem Kriterium des kürzesten Weges gewählt werden muss. Das Phänomen der um den Faktor 4 verringerten Maximalfrequenz wird im Abschnitt 3.2 noch näher erläutert.

Für die analoge Regelungsschaltung, die die LO-Laserfrequenz auf die senderseitige Laserfrequenz abgleichen soll, genügt es, die geschätzte Frequenz pulsweitenmoduliert auszugeben, zu filtern und zu integrieren. Die Analogschaltung muss den Zwischenfrequenzwert auf Null abgleichen. Mit $l = M$, also einer Mittelung über die ohnehin durch den Demultiplexbetrieb parallel zur Verfügung stehenden Sprungzahlen lässt sich die geschätzte Frequenz mit $M + 1$ diskreten Werten darstellen, was für die beschriebene Anwendung vollkommen ausreicht.

Es wurde allerdings im Laborversuch beobachtet, dass die Linienbreite beim geregelten LO-Laser gegenüber dem ungeregelten Fall leicht erhöht ist, die Ausregelung der Zwischenfrequenz wird also womöglich mit erhöhtem Phasenrauschen erkauft. Es ist zu vermuten, dass die Analogschaltung in dieser Hinsicht noch verbesserungsfähig ist, ohne dass die Genauigkeit der Frequenzschätzung erhöht wird.

Eine effektiv nutzbare höhere Genauigkeit würde auch einen höheren Aufwand für die digitale Pulsweitenmodulation voraussetzen. Deshalb ist der im folgenden Abschnitt beschriebene Ansatz für einen Frequenzschätzer mit verbesserter Genauigkeit vorrangig für zukünftige Empfängerkonzepte geeignet, bei denen

wie bei [Leven06] auf eine externe Frequenzregelung verzichtet werden soll.

2.7 Frequenzschätzer mit verbesserter Genauigkeit

Wird eine genauere Frequenzschätzung gefordert, um etwa auf die beschriebene LO-Frequenzregelung zu verzichten und stattdessen die ungeregelte Intradyn-Zwischenfrequenz innerhalb der digitalen Signalverarbeitung zu kompensieren, so empfiehlt es sich, die bei der Sprungzahlbildung auftretende Diskretisierung zu vermeiden und die volle Auflösung der geschätzten Phase auszunutzen. Als Hilfsgröße für die Frequenzschätzung wird zunächst das Phaseninkrement auf dem kürzesten Weg definiert:

$$\zeta(k) := \left(\varphi(k) - \hat{\varphi}(k-1) + \frac{\pi}{4}\right) \bmod \frac{\pi}{2} - \frac{\pi}{4} \tag{2.27}$$

Die Verwendung einer aus der geschätzten Phase $\hat{\varphi}$ abgeleiteten Hilfsgröße ζ bedeutet, dass für die Gesamtdauer der Frequenzschätzung auch noch die Dauer der Phasenrückgewinnung zu berücksichtigen ist. Gravierender als diese Verzögerung ist die Tatsache, dass bei einer Vorverarbeitung der empfangenen Daten durch eine Frequenzkompensation wie bei [Leven06] eine Rückkoppelung erforderlich wäre und die Trägerphasenrückgewinnung nach einer Frequenzkompensation ja gerade keinen linearen Trend mehr aufweisen würde.

Für diese Anwendung sollte ζ deshalb anders definiert werden, nämlich mit aufeinanderfolgenden Werten von ϑ oder ψ an Stelle von $\hat{\varphi}$ in Gl.(2.27). Ein solcher apriorischer Frequenzschätzer arbeitet unabhängig von einem nachfolgenden Phasenschätzer und lässt sich deshalb beispielsweise auch mit einem Asynchronempfänger kombinieren.

In allen Fällen lässt sich die Größe $\zeta(k)$ selbst mit wenigen und in Hardware vermeidbaren Vorfaktoren als Schätzer für die momentane Frequenz verwenden. Arithmetische Mittelung über l Werte $\zeta(k-l)...\zeta(k)$ ergibt einen Schätzer für die Durchschnittsfrequenz im betrachteten Zeitraum.

$$\hat{f}(k,l) = \frac{\langle\zeta\rangle}{2\pi T} \tag{2.28}$$

Die arithmetische Mittelung über sehr viele Werte l liefert allerdings nur dann ein korrektes Ergebnis, wenn die zu schätzende und als stationär angenommene

Zwischenträgerfrequenz klein gegenüber der Abtastfrequenz ist. Für eine Symbolrate von 10 GBaud sind Zwischenfrequenzen von einigen 100 MHz noch unkritisch, die genaue Grenze hängt von weiteren Parametern ab, insbesondere vom Phasenrauschen.

Das Histogramm (also die empirische Häufigkeitsverteilung) einer hinreichend großen Anzahl von $\zeta(k)$ besitzt ein Maximum, dessen Lage die Zwischenträgerfrequenz ebenso repräsentiert wie ein Maximum in einem diskreten Spektrum. Um dieses Maximum zu finden, müsste man statt des arithmetischen Mittels jedoch den Modalwert bilden, was auf eine aufwendige Histogrammberechnung und etliche Vergleichsoperationen hinausläuft.

Da die beschriebene Frequenzschätzung mit dem arithmetischen Mittel aber für alle detektierbaren Frequenzen zumindest das richtige Vorzeichen liefert, lässt es sich für eine Maximumsuche über mehrere Iterationen verwenden. Ein gespeicherter vorheriger Schätzwert $\hat{f}(u-1)$ wird einerseits dazu verwendet das Maximum des ζ-Histogramms in Richtung des Nullpunktes zu verschieben, andererseits dient er nach der Mittelung dazu, den neuen Schätzwert an die richtige Stelle zu verschieben. Ein sinnvoller Starwert ist $\hat{f}(0) = 0$. Der Parameter $h \in]0, 1]$ dient dazu, den iterativen Schätzvorgang zu verlangsamen, um ein Überschwingen zu vermeiden.

$$\hat{f}(u) = \hat{f}(u-1) + \frac{h}{2\pi T} \left\langle \zeta - 2\pi T \hat{f}(u-1) \right\rangle \tag{2.29}$$

Der Frequenzschätzer ist hier rekursiv definiert mit einem erheblich langsamer als k wachsenden Zeitindex u (steht für *update*). Die *Update*-Frequenz braucht nicht konstant zu sein, es erscheint aber zweckmäßig, den Mittelwert immer über eine konstante Anzahl l von verschobenen ζ-Werten zu bilden und dann ohne vermeidbare Verzögerungen eine Aktualisierung nach (2.29) durchzuführen. Ein geeignetes VHDL-Modul für die arithmetische Mittelung wurde im Zusammenhang mit der Polarisationsregelung bereits entwickelt [Samson06, Wörde07].

Der Ansatz, das arithmetische Mittel bzw. ein gewichtetes Mittel über eine Anzahl von Phaseninkrement-Messwerten zu bilden und zur Frequenzschätzung zu verwenden, findet sich auch bei [Kay89]. Der hier vorgestellte Ansatz mit der verschobenen Mittelung ist zugleich ein Verfolgungsalgorithmus (*tracking algorithm*) für eine zeitvariante Frequenz (zur Begriffsabgrenzung von *estimation* und *tracking* vgl. [Quinn]), denn die rekursive Aktualisierung nach (2.29) führt auch bei Veränderung der Frequenz und damit der Lage des Maximums im Histo-

gramm die Frequenzschätzung in der richtigen Richtung nach. Die maximale Geschwindigkeit des Verfolgers wird durch die Wahl von h bestimmt. Umso kleiner h ist, desto stärker wirkt der alte Schätzwert $\hat{f}(u-1)$ als *Bias* für die neue Schätzung.

Kapitel 3

Optimierter Viterbi-Phasenschätzer

3.1 Hardwareeffizienz-Aspekte

Die Elemente des Phasenschätzers nach dem Originalkonzept werden in diesem Abschnitt daraufhin untersucht, wie sie möglichst hardwareeffizient realisiert werden können. Im Laufe der Untersuchung hat sich gezeigt, dass eine Modifikation des Originalkonzeptes (Normierung) sowohl zur Verbesserung des Phasenschätzers als auch zu Hardwareersparnis führt. Originalkonzept und normiertes Konzept sind beide Spezialfälle des Viterbi-Phasenschätzers[Viterbi83].

Die zeitdiskreten elektrischen Eingangswerte $z(k)$ werden von Beginn dieses Kapitels an als wertdiskrete ganze Zahlen betrachtet, die von einem ADC-Paar geliefert werden. Einige der zahlreichen Auswirkungen dieser Umsetzung (Quantisierungsrauschen, Übersteuerung, Kodierungsfehler) werden gesondert behandelt (Begrenzung und Verzerrung). Die Zeitabhängigkeit bzw. logische Abfolge von Größen wie $z = z(k)$ wird aus Gründen der Übersichtlichkeit nur noch notiert, wenn sie in der jeweiligen Formel wichtig ist.

Bei der Umsetzung des Originalkonzeptes taucht neben der Erhebung zur vierten Potenz auch die Aufgabe auf, das Argument $\psi = \text{arc}(z)\text{mod}2\pi$ zu bestimmen. Die Funktionen zur Bestimmung des Argumentes einer komplexen Zahl sollten aus Geschwindigkeitsgründen als Tabellen realisiert werden; Berechnungsverfahren wie z. B. der CORDIC-Algorithmus nach [Volder59] sind demgegenüber langsam und aufwendig [Schwarz99].

Um aus den ADC-Daten das Argument ψ zu gewinnen, verwendet man sinnvollerweise eine Tabelle mit der Arcustangensfunktion. Der Platzbedarf dieser

Tabelle lässt sich auf ein Viertel reduzieren, indem man nach folgendem Schema eine Quadrantenzahl q und die Hilfskoordinaten m, n aus dem Real- und Imaginärteil von z bildet[1]:

sign(Re(z))	sign(Im(z))	q	m	n
+	+	0	$\Re z$	$\Im z$
-	+	1	$\Im z$	$-\Re z$
-	-	2	$-\Re z$	$-\Im z$
+	-	3	$-\Im z$	$\Re z$

Bei Werten auf den Achsen, also mit $\Re = 0$ oder $\Im = 0$ ist die Zuordnung möglichst so, dass $m \neq 0$ ist. Das so gebildete Zahlentripel q, m, n lässt sich wieder in die komplexe Zahl z umrechnen durch

$$z = j^q(m + jn) \tag{3.1}$$

Das Zahlentripel ist also eine Darstellungsform für die komplexe Zahl z, genau wie die kartesische Darstellung, die Polarkoordinatendarstellung oder die Darstellung als Punkt auf der Riemannschen Zahlenkugel. Einige wichtige Eigenschaften dieser Darstellung sind im Anhang zusammengefasst (A.5).

Die Hilfskoordinaten m und n sind nichtnegativ, $m = 0$ tritt außerdem nur bei $z = 0$ auf. Der Tabelleneintrag, der Winkel $\vartheta(m, n)$ aus dem Intervall $[0, \pi/2[$ muss anschließend gemäß (2.12) mit der Quadrantenzahl q verknüpft werden. Die Formel zur Vorab-Berechnung der Tabelleneinträge lautet

$$\vartheta(m, n) = \text{arctan } \frac{n}{m} \tag{3.2}$$

In Hardware werden Quadrantenzahlen q durch 2 Bitstellen MSBQ, LSBQ dargestellt, denen man als Winkel interpretiert die Gewichte π und $\frac{\pi}{2}$ zuordnen kann. Für die Maschinendarstellung von $\psi(k)$ bietet es sich daher an, den Tabelleneintrag von $\vartheta(m, n)$ so zu definieren, dass die Addition aus (2.12) zu einer einfachen Zusammenführung von Leitungen wird. Die Maschinendarstellung von ϑ

[1] Eine weitere Halbierung des Platzbedarfes für die Tabelle wäre durch Ausnutzung der für $m, n > 0$ geltenden Beziehung $\arctan \frac{m}{n} + \arctan \frac{n}{m} = \frac{\pi}{2}$ möglich, ist aber angesichts des zusätzlichen Aufwandes für den Vergleich von m und n und die zusätzliche Umrechnung mit dem Tabelleneintrages nicht sinnvoll. Das Problem mit einer sehr großen Anzahl von Tabelleneinträgen wird beim Polarisationsmultiplex durch vorherige Rundung gelöst.

ist deshalb eine positive Festkommazahl[2] w, deren MSB in logischer Fortsetzung der Quadrantenzahlen der Stellenwert $\frac{\pi}{4}$ zuzuordnen ist. Die Zusammenfassung der Maschinenzahlen q und w zu einer Darstellung des Phasenwinkels ψ aus dem Intervall $[0, 2\pi[$ wird bei der Matlab- und VHDL-Programmierung als Maschinenzahl qw bezeichnet.

Wie beim Winkel ψ ist es auch bei $\hat{\varphi}$ nicht empfehlenswert, die Berechnungen zur Bestimmung des Winkels aus der komplexen Summe jedes Mal einzeln durchzuführen, vielmehr lässt sie sich schnell mit einer Tabelle (*look up table*, LUT) der jeweilige Wert der quadrantenrichtigen Arcustangensfunktion[3] auslesen. Auch die platzsparende Methode der Trennung von Quadrantenzahl und Lagewinkel ist wiederum möglich.

Die Maschinendarstellung von $\hat{\varphi}$ wird als Festkommazahl v bezeichnet. Da ϑ und $\hat{\varphi}$ aus dem gleichen Intervall stammen, ist dem MSB von v wie demjenigen von w ein Stellengewicht von $\frac{\pi}{4}$ zuzuordnen, wenn beim Dekodieren die Differenz $qw - v$ gebildet wird. Die Multiplikation mit dem Faktor $\frac{1}{4}$ aus Formel (2.17) erfolgt implizit durch Stellenzuordnung.

Eine Schwierigkeit bei v ergibt sich aber aus der höheren Auflösung, die die Summe aus (2.17) gegenüber einem einzelnen Wert besitzt. Bei einer exakten Umsetzung der Argument-Funktion müssten entsprechend viele unterschiedliche Tabelleneinträge vorgesehen werden. Es ist daher im Simulationsmodell die Möglichkeit vorgesehen, bei der Verwendung der Summen als Tabellenindex einen Teil der niederwertigsten Bits einfach ungenutzt zu lassen. Dies reduziert zwar den Platzbedarf der Tabelle, aber auch die Genauigkeit der Bestimmung von v.

Es kann sich bei der Summation in der z^4-Ebene exakt der Wert $0+j0$ ergeben, für den das Argument eigentlich nicht definiert ist. Statt der gelegentlich gebräuchlichen aber eigentlich willkürlichen Festlegung $\mathrm{arc}(0) := 0$ ist es in diesem Fall sinnvoller, den Dekodierer ersatzweise den ohnehin verfügbaren ungefilterten Wert w verwenden zu lassen (also $v := w$). Dadurch wird $n_r = q$, also besonders einfach. Gefilterte Alternativen für ein solches nicht definiertes $v(k)$ wären der Vorgänger- oder Nachfolgerwert ($v(k-1), v(k+1)$). Beide Werte sind aber nur aufwendig durch Zugriff auf das entsprechende Nachbarmodul zu gewinnen und

[2] eine feste Zahl von Vorkomma- und eine als Designparameter variable Anzahl von Nachkommastellen erlaubt die Veränderung der Genauigkeit bei dem geforderten gleichbleibendem Gewicht des MSB. Für die Hardware (aber nicht für die VHDL-Beschreibung) besteht zwischen ganzen Zahlen und Festkommazahlen kein Unterschied.

[3] auch als atan2 bezeichnet, weil die Eindeutigkeit durch Verwendung von zwei Funktionsargumenten erreicht wird

könnten im ungünstigsten Fall ebenfalls undefiniert sein. Außerdem könnten in diesem Fall Sprünge nicht erkannt werden, da aus dieser Wahl stets $n_j = 0$ folgt.

Der durch die Winkelzahl w dargestellte Winkel ϑ hat neben seiner ursprünglichen Bedeutung als Lagewinkel im Quadranten nach (2.10) auch noch eine weitere nützliche Bedeutung. Es gilt:

$$\text{arc}(z^4) = 4\psi = 2\pi q + 4\vartheta \tag{3.3}$$

Der erste Term $2\pi q$ kann bei der folgenden Umformung entfallen, weil die Argumentfunktion die Periode 2π besitzt und q eine ganze Zahl ist. Also gilt:

$$\vartheta = \frac{\text{arc}(z^4)\text{mod}2\pi}{4} \tag{3.4}$$

Für die verbesserten Konzepte wird anders als beim Originalkonzept nicht die komplexe Zahl z^4 benötigt, sondern lediglich deren mit (3.3) bereits bekanntes Argument. Deshalb können bei ihnen die aufwendigen komplexen Quadrierungen aus dem Originalkonzept entfallen, so dass die gefundenen Alternativkonzepte nicht nur besser, sondern auch schneller und weniger rechenintensiv sind. Die komplexe Potenzierung, also die innere geschweifte Klammer aus Gl. (2.17), kann durch Ausnutzung dieser Eigenschaft eingespart werden.

3.2 Externer und interner Alias-Effekt

Könnte man alle Rauschquellen und Störeffekte außer dem Phasenrauschen vernachlässigen, so hinge der Lagewinkel nur von der Momentanphase ab und es gälte die Beziehung

$$\vartheta = \left(\frac{\pi}{4} + \varphi\right) \bmod \frac{\pi}{2} \tag{3.5}$$

Beim Intradynempfang ohne Rauschquellen und Störeffekte kann die Zwischenträgerfrequenz zwar nicht als konstant angesehen werden, der Verlauf der Phase müßte aber glatt sein und sich in einem kleinen Betrachtungsintervall näherungsweise linear $mod\frac{\pi}{2}$ beschreiben lassen, was Heterodynempfang mit einem bestimmten Momentanwert der Zwischenträgerfrequenz entspricht.

Für die zuverlässige Funktion der zeitdiskreten Trägerphasenrückgewinnung ist der maximale Momentanwert dieser Zwischenträgerfrequenz entscheidend.

Wird ein kritischer Wert f_{max} überschritten, so kommt es zu Aliaseffekten, d. h. statt der tatsächlichen Frequenz $f > f_{max}$ wird eine falsche Frequenz $f_{alias} < f_{max}$ detektiert.

Aus dem Abtasttheorem ergibt sich unmittelbar, dass für eine korrekte Erfassung des Momentanwertes der Zwischenträgerfrequenz die Differenz zweier aufeinanderfolgender Winkelwerte nicht größer als π sein darf, denn sonst wird durch den Aliaseffekt eine andere Frequenz, evtl. auch mit anderem Vorzeichen detektiert. Das Abtasttheorem lautet

$$f_{max} \leq \frac{1}{2T} \tag{3.6}$$

Da dieser Effekt im Unterschied zu dem nachfolgend behandelten seiner Entstehung nach vor der digitalen Signalverarbeitung liegt, wird er in dieser Arbeit als *externer Aliaseffekt* bezeichnet.

Das abgetastete Signal ist allerdings QPSK-moduliert, so dass zwischen den Winkeln aufeinanderfolgender Werte durch die Modulation schon Sprünge um $\frac{\pi}{2}$ und π auftreten. Es wird bei der Phasenschätzung angenommen, dass derart große Sprünge nicht auf hohe momentane Zwischenfrequenzen zurückzuführen sind.

Diese Annahme wird, bezogen auf den Verlauf der physikalischen Phase, auch Kriterium des kürzesten Weges genannt. Auf die Zwischenträgerfrequenz bezogen bedeutet diese Annahme, dass diese tiefpassbegerenzt sein muss, und der im folgenden erläuterte interne Aliaseffekt beschreibt, was geschieht, wenn sie nicht erfüllt ist: tatsächlich vorhandene hohe Frequenzen werden auf tiefere Aliasfrequenzen abgebildet, was i. A. zu Fehlern bei der Phasenschätzung führt.

Die Zwischenträgerphase wird aus dem abgetasteten Signal z durch eine weitere Operation gewonnen, in der das aufmodulierte Nutzsignal neutralisiert wird. Diese Operation besteht beim Originalkonzept in der Bildung von z^4. Wenn gemäß $z(k) = c(k)e^{j\omega_{IF}kT}$ mit einer konstanten Zwischenträgerfrequenz moduliert wäre, gälte $z^4 \propto e^{j4\omega_{IF}kT}$.

Würde man die Reihenfolge von Abtastung und Potenzierung vertauschen, träten bezogen auf ω_{IF} bereits bei Frequenzen oberhalb eines Viertels von f_{max} gemäß (3.6) Aliaseffekte auf. Bei der tatsächlichen Reihenfolge (Abtastung vor Potenzierung) erhält man aber das gleiche Ergebnis; da diese Berechnungen auch an die Abtatastrate $1/T$ gebunden sind, tritt ein *interner Aliaseffekt* auf. Auch wenn ein Träger bei der vierfachen Zwischenträgerfrequenz physisch gar nicht erzeugt

wird, verhält es sich doch so, als ob dieser noch einmal mit der Periode T abgetastet werden müsste.

Die Anwendung des Abtasttheorems auf eine gedachte Abtastung nach der Potenzierung führt zu der folgenden stärkeren Einschränkung für die maximal detektierbare momentane Zwischnträgerfrequenz:

$$f_{max} \leq \frac{1}{8T} \tag{3.7}$$

Auch die bereits erwähnte Methode, arc(z^4)direkt aus ψ bzw. ϑ zu bestimmen, ist der Einschränkung durch den internen Aliaseffekt unterworfen, dieser ist also unabhängig von der Berechnungsmethode. Wesentlich ist, dass die Berechnung Frequenzanteile erzeugt, die höher sind als die im Ursprungssignal vorhandenen. Solche Operationen sind zur Eliminierung der QPSK-Modulation unvermeidlich.

3.3 Begrenzung und Verzerrung

Für die Gewinnung des Winkels ψ aus dem Eingangswert z mithilfe der Größen q, m, n und ϑ wurde zunächst vorausgesetzt, dass die ADC-Daten von Quantisierungsfehlern abgesehen das analoge elektrische Eingangssignal zum Abtastzeitpunkt kT korrekt abbilden. Tatsächlich ist diese Annahme aber nur im Aussteuerungsbereich der ADCs zulässig, der in der komplexen z-Ebene einer quadratischen Fläche um den Nullpunkt herum entspricht.

Um in der mathematischen Beschreibung unabhängig vom tatsächlichen physikalischen Aussteuerungsberich zu sein (einige 100 mV), wird in dieser Darstellung ein normierte Austeuerungsbereich betrachtet, bei dem jedes z mit $|z| \leq 1$ unverfälscht abgetastet wird.

Ist der Betrag einer der beiden analogen Eingangsspannungen zu groß, so wird für die digitale Signalverarbeitung ein digitaler Wert z auf dem Rand des Aussteuerungsbereiches erzeugt, dessen Amplitudenbetrag zu klein und dessen Phase i. A. verfälscht ist. Das Problem an dieser Phasenverfälschung ist, dass sie bezogen auf den Winkel ϑ einem systematischen Fehler gleichkommt.

Die Viertelkreise stellen analoge komplexe Eingangssignale mit konstanter Amplituden zwischen 1 und 2 in Schritten zu je 0,25 dar. Punkte, die außerhalb des Quadrates links unten liegen (ADC-Aussteuerungsbereich) werden durch die Begrenzung auf den Rand des Quadrates abgebildet, wodurch i. A. die Berechnung

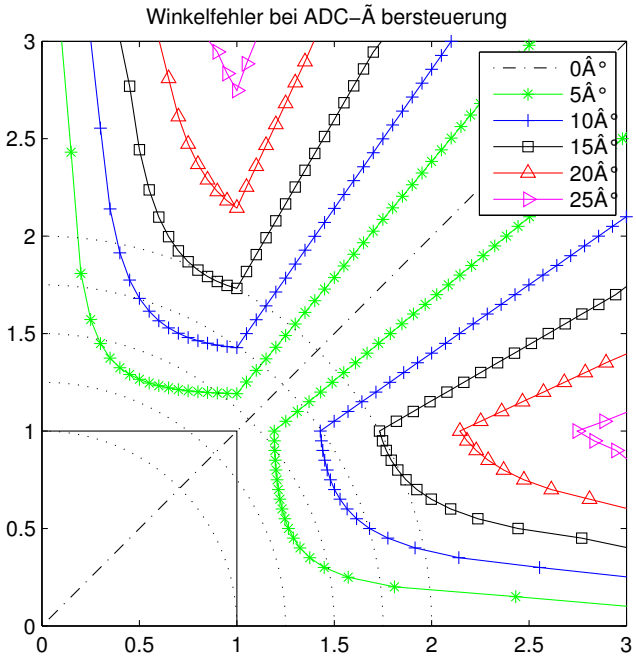

Abbildung 3.1: Winkelfehler bei ADC-Übersteuerung

ihres Argumentes verfälscht wird. Die dargestellten Kurven markieren Mengen von Eingangswerten, bei denen durch die Begrenzung jeweils der gleiche Winkelfehler auftritt. Der Winkelfehler als Kurvenparameter wurde in 5°-Schritten variiert. Die diagonale Strichpunktlinie stellt die Menge aller Eingangswerte dar, die trotz Begrenzung mit einem Winkelfehler von 0° übertragen werden.

Durch Betrachtung der Schnittpunkte von Viertelkreisen und Fehlerkurven lässt sich ablesen, bei welcher Aussteuerung welcher maximale Fehler zu erwarten ist. So beträgt beispielsweis bei einer Eingangsamplitude von 1,25 der begrenzungsbedingte Winkelfehler weniger als 5° (die entsprechenden Kurven werden nicht geschnitten), während er bei einer Eingangsamplitude von 1,5 über 5° betragen kann (Schnitt mit durchgezogenen Kurven).

Die Simulationen zeigen, dass eine gelegentliche leichte Übersteuerung hinsichtlich der Bitfehler unschädlich ist. Wird der Aussteuerungsbereich dagegen nicht ausgenutzt, so wirkt sich der Quantisierungsfehler der ADCs stärker auf die Winkelbestimmung aus, es muss also ein Kompromiss gefunden werden. Die optimale Wahl einer analogen Vorverstärkung entspricht einer Minimierungsaufgabe für den durschnittlichen Winkelfehler.

Die besprochene Begrenzung und Quantisierung ist ein ADC-bedingtes Problem,

das die nachfolgende digitale Signalverarbeitung unabhängig vom verwendeten Phasenschätzer betrifft, weil alle den Winkel ψ verwenden. Die geschilderten Effekte verfälschen aber auch die komplexe Zahl z, aus der nach dem Originalkonzept z^4 gebildet wird.

Alle anderen Konzepte in dieser Arbeit vermeiden die hardwareaufwendige Bildung der Größe z^4, und es soll im Folgenden plausibel gemacht werden, dass ihre Verwendung auch für die Schätzung eher schädlich ist, weil sie nichtlineare Verzerrungen bewirkt. Im folgenden wird erläutert, wie z^4 hardwareeffizient aus z zu gewinnen ist, nämlich anhand der bereits eingeführten Hilfsgrößen q, m, n.

Die Gewinnung des Zwischenfrequenzträgers durch doppelte komplexe Quadrierung der Eingangsdaten und Tiefpassfilterung im digitalen Konzept [Noe04] ist der analogen Lösung in [Noe03] nachempfunden, bei der analoge Spannungssignale miteinander multipliziert und addiert werden sollten. Bei der digitalen Signalverarbeitung liegt z als Paar von ADC-Werten vor, die zur Umsetzung des Originalkonzeptes digital zur 4. Potenz erhoben werden müssen. Dies ist auch für ADCs mit einer Auflösung von 5 Bit relativ aufwendig, vgl. [HPAECOC07]. Verzichtet man auf Rundungen, so besteht z^4 aus zwei 20-Bit-Zahlen. Will man diese Berechung wie bei der Argumentbestimmung durch eine Speichertabelle vermeiden, aber die gleiche Genauigkeit erzielen, wäre der Aufwand allerdings beträchtlich.

Um die digitale Berechnung von z^4 zu erleichtern, kann aber auch auf die zur Winkelbestimmung eingeführten diskreten positiven Hilfskoordinaten m, n zurückgegriffen werden. Die Winkelzahl q braucht wegen $j^{q^4} = 1$ nicht berücksichtigt zu werden, es gilt die folgende Identität:

$$z^4 = (m + jn)^4 \tag{3.8}$$

Ausgehend von einer ADC-Auflösung von 5 Bit sind m und n mit 4 Bit darstellbar, es gibt also $2^4 2^4 = 2^8 = 256$ mögliche Werte für das Paar (m, n) und entsprechend viele Werte für die Bildmenge der Funktion $(m + jn)^4$. Wertemenge und Bildmenge von (3.8) sind in Abb. 3.2 dargestellt. Die Abbildung verdeutlicht die Drehstreckung durch diese Rechenoperation. Die Kurven zwischen den Punkten der Bildmenge verbinden Punkte mit $n = const.$, sind also Bilder horizontaler Gitterlinien. Will man die Berechnung von z^4 durch eine Speichertabelle ersetzten, sollte man Gl. (3.8) ausnutzen, um den Platzbedarf der Tabelle auf ein Viertel zu reduzieren.

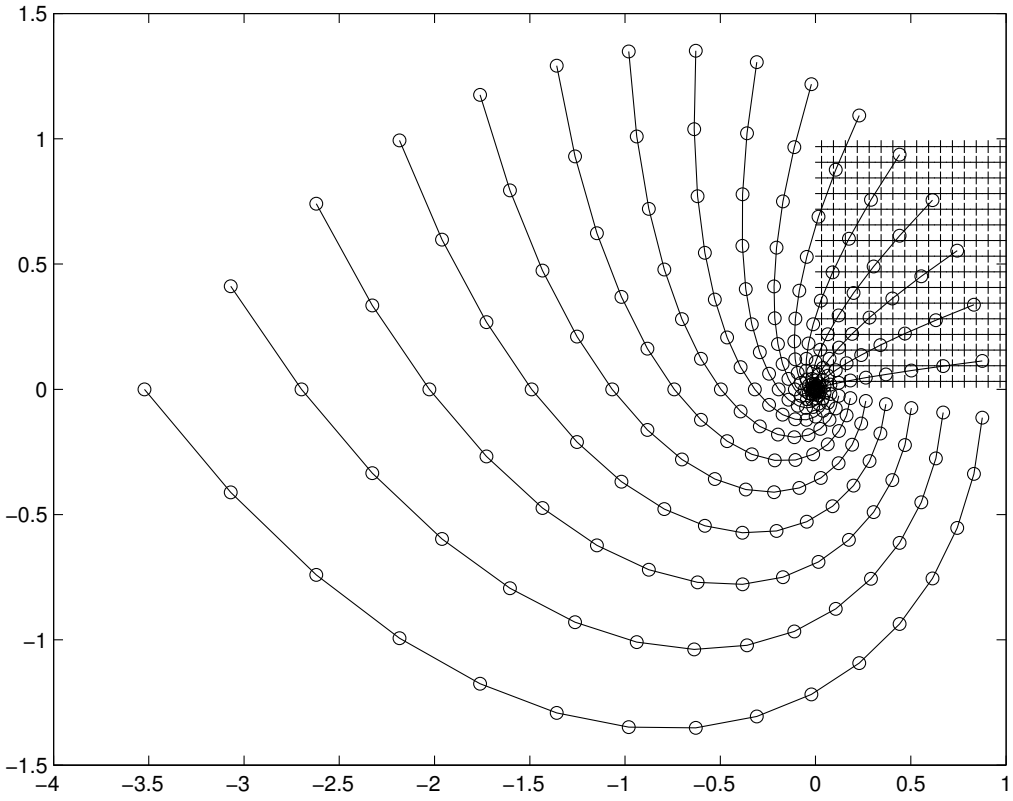

Abbildung 3.2: Wertemenge und Bildmenge von $(m + jn)^4$ für 5-Bit-ADCs

Das quadratische Gitter stellt die bei zwei 5-Bit-ADCs möglichen Werte von m und n dar. Die verzerrte Form des Bildes dieses Gitters durch die Funktion $(m + jn)^4$ macht deutlich, dass durch die Potenzierung zwar die QPSK-Modulation eliminiert wird[4], aber auch mit den ADCs zusammenhängende Effekte wie Quantisierungsrauschen und Begrenzung noch verstärkt werden. Die arithmetische Mittelung des Originalkonzeptes kann aufgefasst werden als eine Schwerpunktbildung unter $2N+1$ auf dem dargestellten "Blatt" liegenden Punkten.

Der Phasenschätzer nach dem Originalkonzept hat die Tendenz, die Werte $\hat{\varphi}$ zum mittleren Wert $\frac{\pi}{4}$ hin zu verfälschen. Dieser systematische Fehler ist verglichen mit dem systematischen Winkelfehler durch leichte Übersteuerung noch stärker ausgeprägt.

[4]genau genommen wird der Einfluß der QPSK-Modulation bereits durch die Verwendung von m und n an Stelle von z eliminiert

3.4 Verbesserung durch Normierung

Das Eingangssignal z unterliegt in seinem Amplitudenbetrag auf Grund des Rauschens und verschiedener Fehlerquellen einer gewissen Schwankung. Da diese Schwankungen für die weitere Verarbeitung keinen Nutzen bringen und durch die Operation z^4 auch noch nichtlinear verzerrt werden, ist es naheliegend, vor der Potenzierung eine Normierung auf den Betrag 1 durchzuführen[5]:

$$z_{normiert} := \frac{z}{|z|} = \exp(j\text{arc}(z)) = e^{j\psi} \tag{3.9}$$

Eine derart auf den Betrag 1 normierte Funktion wird auch als unimodular bezeichnet. Das Resultat der Potenzierung mit 4 liegt bei unimodularen Ausgangswerten wiederum auf dem Einheitskreis. Anstelle von ψ kann für eine Berechnung oder Tabellierung von $z^4_{normiert}$ auch der Lagewinkel ϑ verwendet werden, vgl. (3.3):

$$z^4_{normiert} = \exp(4j\psi) = \exp(4j\vartheta) \tag{3.10}$$

Anstelle des Betrages 1 (Lage auf dem Einheitskreis) kann $z^4_{normiert}$ auch formal mit einer beliebigen konstanten Amplitude $A > 0$ versehen werde. Für die eigentliche Berechnung muss die komplexe Exponentialfunktion berechnet oder tabelliert werden, für das Argument (und damit als Tabellenindex) lässt sich die bereits bestimmte Winkelzahl w verwenden. Realteil c und Imaginärteil s von $z^4_{normiert}$ sind wie folgt zu tabellieren:

$$c(w) \approx A\cos(4\vartheta),\ s(w) \approx A\sin(4\vartheta) \tag{3.11}$$

Die $\approx$-Zeichen weisen dabei darauf hin, dass die trigonometrischen Funktionen nur mit endlicher Genauigkeit tabelliert werden können, so dass Quantisierungsfehler auftreten. Mit den Tabelleneinträgen erhält man das normierte komplexe Signal in kartesischer Darstellung:

$$z^4_{normiert} := c(w) + j \cdot s(w) \tag{3.12}$$

Komplexes Aufsummieren der normierten Werte liefert wie der Originalansatz ohne Normierung eine Vielzahl möglicher Werte für y, zu denen $\hat{\varphi}$ tabelliert

[5] $z \neq 0$ wird vorausgesetzt

werden muss. In Simulationen zeigte sich, dass durch die Normierung die Bitfehlerrate (BER) deutlich gesenkt wird. Dies ist auch leicht zu begründen, denn die Einbeziehung der Amplitude im Originalkonzept verursacht eine wechselnde Gewichtung der unterschiedlichen Summanden, die durch die doppelte Quadrierung auch noch verzerrt ist.

Die Idee, bei der Phasenschätzung die Potenzierung in Polarkoordinaten durchzuführen, findet sich bereits bei [Viterbi83]. Dort werden drei verschiedene Phasenschätzer miteinander verglichen[6]:

$$\hat{\varphi}(k) = \frac{1}{4}\text{arc}\left(\sum_{n=k-N}^{k+N} |z(n)|^p \, e^{j4\psi(n)}\right) \quad p \in \{0, 2, 4\} \tag{3.13}$$

Für den Fall $p = 0$ ist der Phasenschätzer nach [Viterbi83] also mit dem hier hergeleiteten normierten Konzept identisch. Für $p = 4$ erhält man das Originalkonzept in Polarkoordinatendarstellung, was in diesem Fall dem Aufwand lediglich vergrößert, weil der Betrag $|z(n)|$ gebildet, potenziert und mit dem Phasor $e^{j4\psi(n)}$ multipliziert werden muss, obgleich er wie dargelegt eigentlich keine für die Phasenschätzung nützliche Information liefern kann[7].

Auch der Viterbi-Phasenschätzer mit $p = 2$ wurde auf Grund des hohen Aufwandes und zweifelhaften Nutzens nicht näher untersucht. Stattdessen wurde ein weiterer Ansatz verfolgt, bei dem das einfache arithmetische Mittel in Gl. (2.17) und (3.13) durch eine verbesserte Tiefpassfilterung ersetzt wird.

3.5 Verbesserung durch Gewichtung

Die arithmetische Mittelung über $2N + 1$ Werte wird im Originalkonzept und bei [Viterbi83] eingesetzt, um aus der Zeitreihe der bereits durch Potenzierung von der QPSK-Modulation befreiten komplexen Werte das mittelwertfreie additive Rauschen herauszufiltern. Durch die symmetrische Einbeziehung von Vorgängern und Nachfolgern wird auch der lineare Trend der Phase ausgeglichen, den eine nichtverschwindende Zwischenträgerfrequenz verursacht. Um diesen Vorteil bei einer Optimierung nicht zu verlieren, sollte das Prinzip der Symmetrie beibehalten werden.

[6] der für die Argumentbildung bedeutungslose Vorfaktor $\frac{1}{2N+1}$ der Summe wurde weggelassen und insgesamt die Nomenklatur angepasst

[7] Nur bei nichtlinearem Phasenrauschen, welches in dieser Arbeit aber vernachlässigt wird, könnte die Betrachtung der Amplitude bei der Phasenschätzung nützlich sein.

Im Originalkonzept wurde vorausgesetzt, dass $2N + 1$, also eine ungerade Anzahl, komplexer Summanden mit gleicher Gewichtung gemittelt werden. Die für die technische Realisierung ungünstige Division mit $2N + 1$ kann bei der Mittelwertbildung jedoch problemlos durch Division mit einer Zweierpotenz ersetzt oder ganz weggelassen werden, weil sich durch diese Operationen das Argument des Mittelwertes bzw. der komplexen Summe nicht ändert.

In der statistischen Literatur sind für Zeitreihen zwei Mittelwertdefinitionen gebräuchlich: der gerade und der ungerade. Der ungerade entspricht dem Originalkonzept und gewichtet alle Summanden gleich, der gerade gewichtet dagegen die Randwerte nur halb so stark wie die übrigen.

$$M_{ungerade}(x_n) = \frac{1}{2N+1}\left(x_{n-N} + x_{n-N+1} + ... + x_{n+N}\right) \tag{3.14}$$

$$M_{gerade}(x_n) = \frac{1}{2N}\left(\frac{x_{n-N}}{2} + x_{n-N+1} + ... + x_{n+N-1} + \frac{x_{n+N}}{2}\right) \tag{3.15}$$

Experimentell wurde auch noch eine dritte Mittelwertbildung (ohne eingeführte Bezeichnung in der Literatur) gefunden, deren Simulationsergebnis besser war als die bekannten Mittelwertfomeln: Sie gewichtet den mittleren Wert doppelt im Verhältnis zu allen anderen Werten.

$$M_{Mittenverdopplung}(x_n) = \frac{1}{2N+2}\left(x_{n-N} + x_{n-N+1} + ... + 2x_n + ... + x_{n+N}\right) \tag{3.16}$$

Diese Verdoppelung ist in der Praxis sehr einfach: entweder man konstruiert einen Addierebaum für $2N + 2$ Werte und speist den mittleren Wert doppelt ein, oder man multipliziert den Mittelwert mit dem Faktor 2, was für Festkommazahlen sehr leicht realisierbar ist. Wie bei geradem und ungeradem Mittelwert ist auch bei Verdoppelung des mittleren Wertes die Gewichtung insgesamt symmetrisch.

Die doppelte Gewichtung des Zentralwertes liefert im Simulationsvergleich gegenüber geradem und ungeradem Mittelwert das beste Ergebnis. Damit ist aber noch nicht gesagt, dass diese Mittelwertbildung nicht noch weiter verbessert werden kann. Ein allgemeinerer Ansatz als die beschriebenen Mittelwerte ist ein Digitalfilter mit frei wählbaren Gewichtungen der Summanden

Das arithmetische Mittel wird verallgemeinert durch eine Gewichtung der Sum-

manden mit beliebigen reellen positiven Faktoren g_n. Symmetrie um den zeitlichen Mittelpunkt bei k zur Unterdrückung des linear steigenden oder fallenden Teils führt auf die Form

$$y(k) = g_0 e^{j4\vartheta(k)} + \sum_{n=1}^{N} g_n (e^{j4\vartheta(k+n)} + e^{j4\vartheta(k-n)}) \tag{3.17}$$

Neben der Symmetrieforderung erscheint es plausibel, bei einer ungleichen Gewichtung der Summanden diejenigen Summanden besonders stark zu gewichten, die zeitlich nahe am mittleren Wert (Zeitpunkt kT) sind. Dem entspricht die Forderung

$$g_0 \geq g_1 \geq g_2 \geq ... \tag{3.18}$$

O.B.d.A. kann man das mittlere Gewicht auf den Wert $g_0 = 1$ festlegen. Durch eine solche symmetrisch um g_0 fallende Gewichtung können die beim einfachen arithmetischen Mittel bestehenden Vorzüge eines großen Wertes N (gute Glättung dank vieler einbezogener Werte) und eines kleine Wertes N (schnelle Änderung der physikalischen Phase sind verfolgbar) miteinander kombiniert werden.

Die gewichtete Mittelung kann auch interpretiert werden als Faltung mit der endlichen Impulsantwort eines entsprechenden diskreten Filters. Den Vorzug einer solchen gewichteten Mittelung gegenüber der einfachen arithmetischen Mittelung kann man sich gut an der Sprungantwort[8] dieses Filters verdeutlichen: während bei der arithmetischen Mittelung der Anstieg der Ausgangsgröße mit konstanter Geschwindigkeit erfolgt (Treppenfunktion), ist bei Gewichtung gemäß (3.18) der Anstieg am Anfang und Ende der Sprungantwort eher flach, während er in der Mitte am steilsten ist. Bei einem zeitkontinuierlichen Filter würde das einer S-förmigen Kurve entsprechen.

Das normiert-gewichtete Konzept ist allerdings recht aufwendig zu realisieren. Insbesondere die komplexe Multiplikation mit den Gewichten g_n würde gegenüber der einfachen Normierung einen erheblichen Mehraufwand bedeuten. Deshalb wurde nach einer Optimierung in Schritten zu $0,05 = \frac{1}{20}$ für eine spätere Hardwareumsetzung als zusätzliches praktisches Kriterium die ganzzahlige Darstellbarkeit der g_n mit 3 Bit gefordert. Bezogen auf $g_0 = 1$ können die Gewichte

[8] Man beachte, dass das Eingangssignal komplex ist. Die folgende Aussage gilt sowohl für einen Sprung von Realtteil und/oder Imaginärteil von $z(k)$ als auch für einen Phasensprung dieses Eingangssignals.

also 7 verschiedene Werte $\frac{1}{8}...\frac{7}{8}$ annehmen.

Mit dieser Vorgabe wurde vergleichende Monte-Carlo-Simulationen zur Bestimmung der besten Werte der g_n durchgeführt. Die Ergebnisse für $N = 3...7$ sind in der folgenden Tabelle wiedergegeben:

N	$g_n = \frac{i_n}{20}$	$g_n = \frac{i_n}{8}$
3	$0,5; 0,4; 0,35$	$\frac{4}{8}; \frac{3}{8}; \frac{2}{8}$
4	$0,45; 0,35; 0,25; 0,2$	$\frac{4}{8}; \frac{3}{8}; \frac{2}{8}; \frac{1}{8}$
5	$0,5; 0,35; 0,25; 0,2; 0,15$	$\frac{4}{8}; \frac{3}{8}; \frac{2}{8}; \frac{1}{8}; \frac{1}{8}$
6	$0,45; 0,35; 0,3; 0,2; 0,15; 0,1$	$\frac{4}{8}; \frac{3}{8}; \frac{2}{8}; \frac{2}{8}; \frac{1}{8}; \frac{1}{8}$
7	nicht optimiert	$\frac{4}{8}; \frac{3}{8}; \frac{2}{8}; \frac{2}{8}; \frac{1}{8}; \frac{1}{8}; \frac{1}{8}$

Mit i_n sind jeweils allgemein diejenigen natürlichen Zahlen bezeichnet, die sich als Ergebnis der Optimierung und vorgegebenen Rundung ergeben haben. Die Optimierungsergebnisse lassen sich annähernd verallgemeinern durch die Funktion $g(n)$:

$$g_n \approx g(n) = \frac{1}{1+n} \tag{3.19}$$

Für eine Beschreibung der gewichteten Mittelung als FIR-Filter (*finite impulse response*) gilt $g_n = 0$ für $n > N$. Theoretisch wäre es gut, möglichst viele Abtastwerte in die Filterung einzubeziehen, um das additive Rauschen möglichst gut zu unterdrücken. Praktisch ist der mögliche Aufwand aber begrenzt, und da ein großer Wert für N auch die maximal detektierbare momentane Zwischenfrequenz begrenzt, gibt es ein endliches Optimum für die Wahl von N. Bei den oben tabellierten Beispielen lieferte die Variante mit $N = 5$ das beste Simulationsergebnis [HofCOTA].

Abb. 3.3 stellt für drei Phasenschätzer (original, normiert, normiert-gewichtet) mit $N = 5$ die BER-Kurven dar. Außerdem sind die BER-Kurven für den Asynchronempfänger ($N = 0$) und für einen idealen Phasenschätzer dargestellt. Der ideale Phasenschätzer verwendet den nur im Simulationsprogramm verfügbaren ursprünglichen Verlauf der Zwischenträgerphase (mit Phasenrauschen, aber ohne AWGN) und ist somit nicht realisierbar, aber sein Verhalten ist realistischer als das eines rein theoretischen Phasenschätzers, welcher in [HofCOTA] ebenfalls behandelt wurde. Das gängige theoretische Modell berücksichtigt insbesondere nicht die Fehlerverdoppelung bei differenzieller Kodierung.

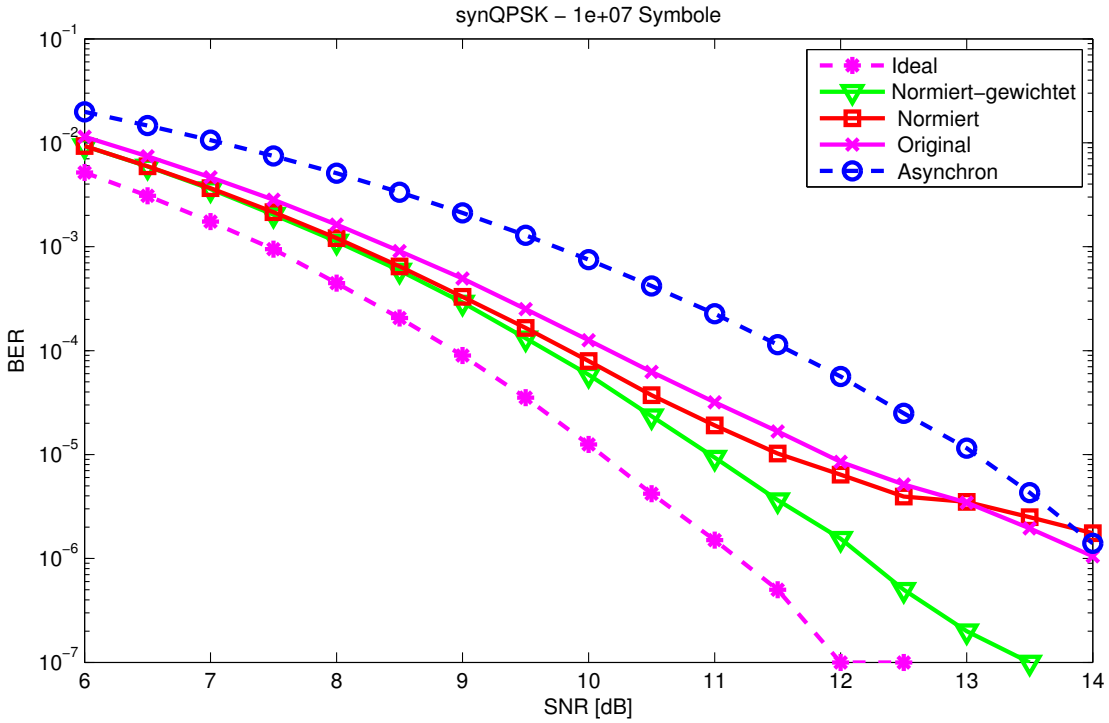

Abbildung 3.3: BER-Kurven für Viterbi-Phasenschätzer

Die gefundene Formel (3.19) für den zeitdiskreten Fall lässt sich in eine unendliche nichtkausale Sprungantwort übersetzen:

$$g(t) = \frac{1}{1 + |t|} \tag{3.20}$$

In der Literatur, z.B. [IpKahn07, Taylor05] zu einer Phasenschätzung mit gewichteter Mittelung wird dagegen das Wiener-Filter [Wiener] als theoretisches Optimum angeführt. Die Theorie zum Wienerfilter, das eine optimale Rauschunterdrückung zu liefern verspricht, wurde schon in den 1940er Jahren entwickelt. Das Nutzsignal (also hier der IF-Phasor) und das additive Rauschen werden dabei als stochastische Prozesse mit bekannter Spektralverteilung oder bekannter Autokorrelation und Kreuzkorrelation vorausgesetzt.

3.6 Zusammenfassung

Das in [Noe04] veröffentlichte Konzept wurde als hardwarenahes Matlab-Modell implementiert, wobei zunächst einige Modifikationen erforderlich waren, die

zu Beginn des Kapitels erklärt wurden. Nach Simulationen und Analyse der Schwachstellen des Originalkonzeptes wurde in zwei Schritten optimiert, wobei theoretische Analyse und Simulation miteinander verzahnt wurden.

Auf diese Weise wurde ein erheblich leistungsfähigeres alternatives Konzept entwickelt, bei dem die Phasenschätzung grundsätzlich verbessert wurde: von den Messdaten (komplexe Zahlen in kartesischen Koordinaten) wurde nur die Phaseninformation verwendet, die ohnehin für die Dekodierung mithilfe einer Tabelle gewonnen werden muss. Der Betrag (Amplitude, Radius) der Messdaten wurde dagegen aus der Berechnung eliminiert. Dieses Vorgehen ist allgemein dadurch gerechtfertigt, dass bei einem Phasenmodulationsverfahren die nur durch Rauschen bzw. Messfehler veränderte Amplitude keine nützliche Information enthalten kann.

Die Normierung und die durch Polarkoordinaten erleichterte Potenzierung finden sich auch schon als eine mögliche Phasenschätzervariante in der vielzitierten älteren Veröffentlichung [Viterbi83]. Leider wird bei vielen aktuellen Veröffentlichungen aber nicht erklärt, welche Variante des Viterbi-Phasenschätzers realisiert wurde.

Die Elimination der Amplitudeninformation und Erzeugung einer unimodularen Ersatzfunktion für das Eingangssignal verursacht zwar zunächst einigen Zusatzaufwand (Tabellen für die komplexe Exponentialfunktion), andererseits wird die Berechnung aber insgesamt leichter, weil die Erhebung des Eingangssignals zur 4. Potenz nicht mit komplexen Multiplikationen realisiert werden muss. Stattdessen kann einfach der Faktor 4 im Exponenten bei der Tabellierung der komplexen Exponentialfunktion berücksichtigt werden. Der eigentliche Zweck der doppelten Quadrierung, nämlich die Elimination der QPSK-Modulation, wird in der praktischen Umsetzung sehr einfach durch Weglassen der Quadrantenzahl q erreicht.

Die auch schon bei Untersuchungen am Originalkonzept als nützlich erkannte unterschiedliche Gewichtung der einzelnen Summanden wurde auf das normierte Konzept übertragen. Es ist erwiesenermaßen von Vorteil, dass hier ausschließlich die vorgegebenen Gewichtsfaktoren und nicht mehr der ursprüngliche oder gar der durch Potenzierung verzerrte Betrag den Einfluss eines einzelnen Messwertes auf die Summe und damit auf die geschätzte Phase bestimmen.

Einige einleuchtende Grundregeln für die Wahl der Gewichtsfaktoren wurden vorgestellt, konkrete durch Monte-Carlo-Simulation gewonnene Werte doku-

mentiert und Bezüge zu einer allgemeinen Theorie (Wiener-Filter) aufgezeigt. Die VHDL-Umsetzung des durch Normierung und Gewichtung gekennzeichneten verbsserten Konzeptes wurde im Rahmen einer Studienarbeit geleistet [Romoth07]. Gegenüber dem im folgenden Kapitel vorgestellten besonders hardwareeffizienten Phasenschätzer war das Konzept allerdings zu aufwendig für eine praktische Umsetzung.

Kapitel 4

Direkte Phasenschätzung

4.1 Grundidee und allgemeine Beschreibung

Betrachtet man das im vorigen Kapitel beschriebene optimierte Phasenschätzerkonzept abstrakt als ein System, das in Abhängigkeit von Eingangsgrößen ϑ eine Ausgangsgröße $\hat{\varphi}$ erzeugen soll, so wird deutlich, dass es sich trotz der enthaltenen nichtlinearen und komplexen Funktionen letzlich um eine Filterfunktion handelt. Deren Eingangs- und Ausgangsgrößen sind reell und in ihrem Wertebereich beschränkt, da es sich um Winkelgrößen handelt.

In Maschinendarstellung mit beschränkter Auflösung gibt es sogar nur endlich viele mögliche Eingangskombinationen, und da das System deterministisch und gedächtnisfrei ist, ließe sich die komplette Filterfunktion sogar tabellieren, der Aufwand wäre allerdings beträchtlich und nicht hardwareeffizient. Immerhin handelt es sich um die skalare Funktion einer mehrdimensionalen Eingangsgröße, also als Tabelle um ein Feld (*Array*) mit der Dimension $2N + 1$.

Andererseits ist auch die im vorigen Kapitel beschriebene Realisierung, die sich aus der theoretischen Beschreibung mehr oder weniger direkt ergibt, in der Praxis kompliziert und durch die Verkettung von Tabellen und die komplexe gewichtete Addition relativ langsam. Eine sinnvolle Alternative wäre deshalb eine einfacher realisierbare heuristische Ersatzfunktion. Das Verhalten des Filters braucht dabei nicht in allen Details nachgeahmt zu werden, es handelt sich nur um einen Ersatz, dessen Qualität sich in der Praxis durch BER-Vergleich mit der bekannten zu ersetzenden Funktion erweisen muss.

Da die in diesem Kapitel vorgestellten Ersatzfunktionen sich zwar grundsätz-

lich als Übertragungsfunktion für ein komplexes Signal darstellen lassen, diese Beschreibung aber sehr unübersichtlich wäre (Nichtlinearitäten, keine geschlossenen Ausdrücke, Fallunterscheidungen) und sie eigentlich eher den Charakter von Suchalgorithmen haben, wird in diesem Kapitel gegenüber den Begriffen Winkelfilter oder Phasenrückgewinnung der Begriff Winkelschätzer (*Phase estimator*) bevorzugt. Dadurch soll auch betont werden, dass der Schätzer auch noch bei eigentlich unbrauchbaren Eingangsdaten einen brauchbaren Ausgangswert liefern muss.

Da bereits im Originalkonzept vom komplexen tiefpassgefilterten Signal y nur das Argument in Form des Winkels $\hat{\varphi}$ verwendet wird und bei den bisher vorgestellten verbessertern Filtern eingangsseitig nur der Lagewinkel ϑ (in Hardware: die Winkelzahl w) verwendet wird, ist es naheliegend, eine Filterfunktion F_W direkt auf der Basis der mit w bereits gegebenen Argumente von z^4 zu konstruieren. Wie beim komplexen Filter können zur Aufwandsbeschränkung nur $2N+1$ Werte aus der direkten zeitlichen Umgebung verwendet werden, wobei wieder die symmetrische nichtkausale Notation verwendet wird:

$$\hat{\varphi}(k) = F_W(\vartheta(k-N), ..., \vartheta(k+N)) \tag{4.1}$$

Eine solche Winkelfilterfunktion[1] F_W kann ganz anders aufgebaut sein als eines der im vorigen Kapitel beschriebenen kartesischen Filter[2]. Um ähnlich gute Ergebnisse wie bei den kartesischen Filtern zu liefern, sollte deren Verhalten aber möglichst gut nachgebildet werden.

Im Verlauf dieses Kapitels werden zwei verschiedene Winkelschätzerkonzepte vorgestellt, von denen das zweite sich nicht nur in Simulationen sondern mittlerweile auch in praktischen Experimenten [PfauCOTA] bewährt hat. Es zeichnet sich gegenüber dem normiert-gewichteten Konzept aus dem vorigen Kapitel durch besonders hohe Hardwareeffizienz aus [HofCOTA].

Eine reelle Phasenschätzerfunktion nach Gl. (4.1) ist immerhin noch eine Abbildung $\mathbb{R}^{2N+1} \rightarrow \mathbb{R}$. Besonders günstig für die Realisierung (also hardwareeefizient)

[1] d.h. Funktion eines Filters, dessen Ein- und Ausgangsgrößen Winkel sind, im Gegensatz zu kartesischen Filtern, die mit komplexen Zahlen in kartesischen Koordinaten arbeiten und den ebenfalls komplexen Viterbi-Phasenschätzern, die intern auch noch die Polardarstellung verwenden

[2] auch wenn die Phasenschätzer nach [Viterbi83] Polarkoordinaten verwenden, so erfolgt doch die komplexe Mittelwertbildung dort in kartesischen Koordinaten, weshalb sich das Attribut 'kartesisch' zur Unterscheidung von den in diesem Kapitel behandelten direkten oder winkelbasierten Phasenschätzern besser eignet als z. B. 'komplex'.

wäre eine Ersatzfunktion, die als Verkettung mehrerer gleichartiger Funktionen $\mathbb{R}^2 \rightarrow \mathbb{R}$ darstellbar ist, deren Zwischenergebnisse auch noch mehrfach benutzbar wären, analog zu den Zwischensummen bei der Realisierung der kartesischen Filter. Dies ist die Grundidee des zweiten Konzeptes, das deshalb auch als verteilte Mittelwertbildung bezeichnet wird.

Um für die verteilte Mittelwertbildung eine brauchbare Teilfunktion[3] zu finden, ist es hilfreich, die entsprechende Berechnung bei den komplexen Teilsummen aus dem normiert-gewichteten Konzept genauer zu analysieren. Betrachtet wird daher im Anhang als Teilproblem die Summe zweier unimodularer komplexer Zahlen mit vorgegebenen reellen Gewichtungen. Direkt wird das Ergebnis in Abschnitt 4.3 hergeleitet. Anschließend werden die Eigenschaften einer geeigneten Baumstruktur zur verteilten Mittelwertberechnung behandelt.

4.2 Drehwinkelfilter

Ein Ansatz[4], nach dem über $2N+1$ Winkelwerte gemittelt werden kann, basiert auf der Bestimmung eines geeigneten Verdrehungswinkels α, um den alle Lagewinkel vor der Mittelung geändert werden:

$$\vartheta'(k) := (\vartheta(k) - \alpha) \text{mod } \frac{\pi}{2} \tag{4.2}$$

Diese Verdrehung leistet dasselbe wie später die Einbeziehung der Differenz in (4.6); das einfache arithmetische Mittel kann somit wie folgt gebildet werden:

$$\mu' = \frac{1}{2N+1} \sum_{k=1}^{K} \vartheta'(k) \tag{4.3}$$

Nach dieser normalen Mittelwertbildung muß nur noch die Verdrehung wieder rückgängig gemacht werden :

$$\mu = (\mu' + \alpha) \text{mod } \frac{\pi}{2} \tag{4.4}$$

[3]die entsprechende Hardware wird im folgenden als Elementarzelle bezeichnet. Die Elementarzelle ist also zunächst einfach eine digitale Teilschaltung, die aus zwei Winkelwerten einen neuen Wert erzeugt, und der Phasenschätzer insgesamt ist eine Baumstruktur mit Elementarzellen als Knoten.

[4]Das Drehwinkelfilter-Konzept basiert auf einem Vorschlag von Prof. Noe, wurde aber verallgemeinert und in verschiedenen Varianten simuliert.

Das Verfahren verursacht allerdings hohen Aufwand durch die Subtraktion und Addition des Verdrehungswinkels α; je mehr verschiedene abgestufte Verschiebungsmöglichkeiten vorgesehen werden, umso aufwendiger ist die Implementierung. Der geringstmögliche aber noch sinnvolle Aufwand für die Verschiebung besteht in der Auswahl eines Quadranten[5], also der Beschränkung auf vier verschiedene Verschiebungsmöglichkeiten. Bleibt ein Quadrant frei, so wird dieser durch die gemeinsame Verschiebung zum Quadranten mit der höchsten Nummer gemacht. Die verschobenen Werte und der aus ihnen gebildete Mittelwert liegen somit in den ersten drei Quadranten und damit unter Berücksichtigung der Verdrehung an der richtigen Stelle. Dieses Verfahren setzt allerdings voraus, dass ein ganzer Quadrant unbesetzt geblieben ist, sonst kann der Algorithmus keine passende Verdrehung für eine korrekte Mittelwertbildung finden. Dadurch ist diese einfachste Realisierung des Konzeptes nicht praxistauglich.

Die höchste praktikable Auflösung von α ist dieselbe wie die von ϑ. Ein Suchalgorithmus ordnet α den Wert direkt oberhalb der "größten Lücke"[6] zu. Dieses Konzept funktioniert, hat aber einen enorm hohen Aufwand zur Folge, denn in der Grundmenge von $2N + 1$ Werten müssen zur Bestimmung von Maximum und Minimum je nach Sortieralgorithmus bis zu $(2N + 1)!$ Vergleiche durchgeführt werden. Das Endergebnis μ ist allerdings auch hier nur dann korrekt, wenn sich tatsächlich ein geeigneter Winkel α bestimmen lässt. Schon mit $N = 1$ gibt es problematische Fälle, in denen der Mittelwert nicht korrekt gebildet werden kann.

Ein weiterer Nachteil dieses Verfahrens besteht darin, dass die Verschiebung allgemein in jedem Modul[7] mit einem anderen Wert als beim Nachbarmodul erfolgen muss, von mehreren Modulen gemeinsam nutzbare Zwischenergebnisse wie bei den anderen Konzepten kommen nicht vor.

Im vorigen Kapitel wurde dargelegt, dass es vorteilhaft ist, die zeitlich mittleren Werte stärker zu gewichteten als die vom zu schätzenden Wert zeitlich entfernteren. Eine solche stärkere Gewichtung der zeitlich mittleren Werte lässt sich beim Drehwinkelfilter-Konzept nur durch mehrfaches Einspeisen dieser Winkelwerte erreichen. Die Zuverlässigkeit des Verfahrens wird dadurch nicht verändert, weil

[5] 'Quadrant' bezieht sich hier auf die $z^4_{normiert} = e^{j4\vartheta}$-Ebene; im Verfahren nach [**?**] ist die Bestimmung eines 'exclude-Quadranten' also gleichbedeutend mit der Festlegung $\alpha = \frac{x\pi}{8}$mit $x \in \{0, 1, 2, 3\}$ für Gl. (4.2)

[6] da $\vartheta \in [0, \frac{\pi}{2}[$nach 3.3 das Argument von z^4repräsentiert, bezieht sich das auf die größte Lagewinkeldifferenz mod $\frac{\pi}{2}$.

[7] Ein Modul bildet im Demux-Betrieb u.a. den jeweiligen Filterwert, vgl. [Noe04]

die eingespeisten Werte einander gleich sind, aber der Aufwand steigt.

Exemplarisch wurde Doppeleinspeisung des mittleren Wertes simuliert, sie verbessert auch erwartungsgemäß leicht das Ergebnis. Weitere Abstufungen der Gewichtung wären allerdings nur mit beträchtlichem Aufwand möglich, weil die mehrfach eingespeisten Werte zumindest bei der Mittelung auch entsprechend mehrfach in die Berechnung einfließen müssen. Außerdem verbessern sie bei diesem Konzept das Ergebnis nicht wesentlich, weil dessen Hauptproblem die Anfälligkeit gegenüber Ausreißern ist, die das Auffinden der richtigen Lücke erschweren können. Insgesamt wurde dieses Konzept wegen des hohen Aufwandes und der nicht überzeugenden Ergebnisse nicht weiter verfolgt.

4.3 Verteilte Mittelwertbildung

Gegeben seien statt $2N + 1$ zunächst nur zwei Lagewinkel[8] α und β, die zeitlich nicht notwendigerweise direkt benachbart sein müssen. Aus ihnen soll ein Schätzwert μ gebildet werden, den man als mittleren Wert einer linearen Phaseninterpolation zwischen zwei Messpunkten interpretieren kann. Es wird gefordert, dass μ wiederum ein Lagewinkel ist, also in $[0, \frac{\pi}{2}[$ liegt. Um einen solchen Schätzwert zu finden, bietet sich zunächst das arithmetische Mittel an:

$$\mu = \frac{1}{2}(\alpha + \beta) \tag{4.5}$$

Dies ist zwar stets ein Winkel aus dem Intervall $[0, \frac{\pi}{2}[$, aber nicht immer der korrekte Mittelwert im Sinne der Aufgabe, denn der Schätzwert sollte immer auf dem kürzesten Weg $mod\frac{\pi}{2}$ gebildet werden. Beträgt der Differenzbetrag der beiden Winkel mehr als $\pi/4$, so muss bei μ gemäss 4.5 für die korrekte Mittelung noch der Wert $\pi/4$ addiert bzw. subtrahiert werden, was in Hardware leicht und einheitlich durch eine Addition $mod\frac{\pi}{2}$ erfolgen kann. Man benötigt für die Berechnung als Hilfsgrößen die Summe $\sigma = \alpha + \beta$ und die Differenz $\delta = \alpha - \beta$ der Eingangswinkel und erhält den korrekten Mittelwert μ durch Fallunterscheidung:

[8] also ein Winkel ϑ aus $[0, \pi/2[$, zur Unterscheidung umbenannt in α und β. Die beiden Lagewinkel sind bei der Schätzung gleich stark zu gewichten.

$$\mu = \begin{cases} \sigma/2, & |\delta| < \pi/4 \\ \sigma/2 + \pi/4, & |\delta| > \pi/4, \quad \sigma < \pi/2 \\ \sigma/2 - \pi/4, & |\delta| > \pi/4, \quad \sigma \geq \pi/2 \end{cases} \tag{4.6}$$

Mathematisch gesehen herrscht Nichtentscheidbarkeit bei $|\delta| = \pi/4$ (zwei gleichlange Wege, kein kürzester). In der Praxis, also bei Realisierung in Soft- oder Hardware muss diese Entscheidung aber getroffen werden. Von der tatsächlichen Behandlung dieses Grenzfalles hängt es ab, ob für diesen Schätzer die Vertauschbarkeit von α und β gilt, also die Kommutativität $\mu(\alpha, \beta) = \mu(\beta, \alpha)$.

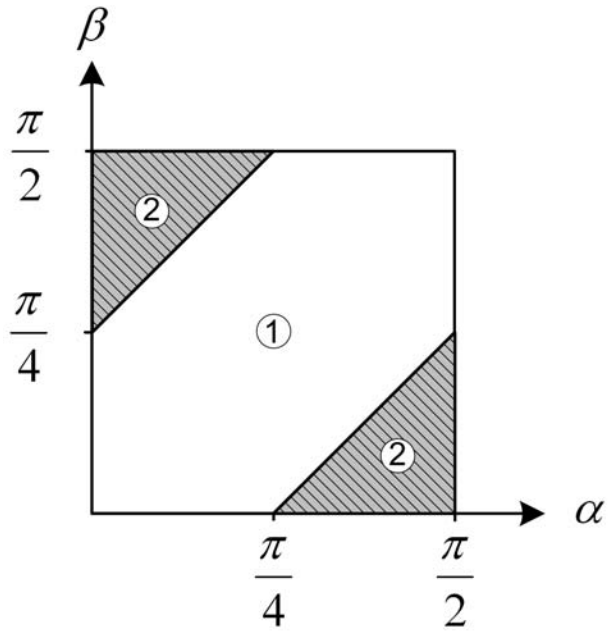

Abbildung 4.1: Geltungsbereiche der Mittelwertformeln

Abb. 4.1 zeigt die Geltungsbereiche der verschieden Mittelwertformeln in Abhängigkeit von den Eingangsgrößen α und β. In Bereich 1 gilt $\mu = \frac{\alpha+\beta}{2}$, in den schraffierten und mit 2 markierten Bereichen muss dagegen $\frac{\pi}{4}$ dazuaddiert bzw. abgezogen werden. Ist der Bereich 1 gegenüber den Bereichen 2 einheitlich ein offenes oder geschlossenes Gebiet, so ist die Mittelwertbildung kommutativ. Dies ist einsichtig, weil beim Grenzfall $|\delta| = \pi/4$ einheitlich für $\delta = +\pi/4$ und $\delta = -\pi/4$ entschieden wird, so dass sich nichts ändert, wenn man δ durch $\delta' = \beta - \alpha$ ersetzt.

Es folgt noch eine geschlossene Darstellung der Mittelwertformel, bei der die Festlegung gilt, dass Bereich 1 in Abb. 4.1 beidseitig gegenüber den Bereichen 2 geschlossen ist:

$$\mu = \left\{ \frac{\sigma}{2} + \frac{\pi}{4} \left\lceil \frac{|\delta|}{\pi/4} \right\rceil \right\} \bmod \frac{\pi}{2} \tag{4.7}$$

Diese Formel für μ ist gegenüber Gl.(4.6) auch bei $|\delta| = \frac{\pi}{4}$ definiert.

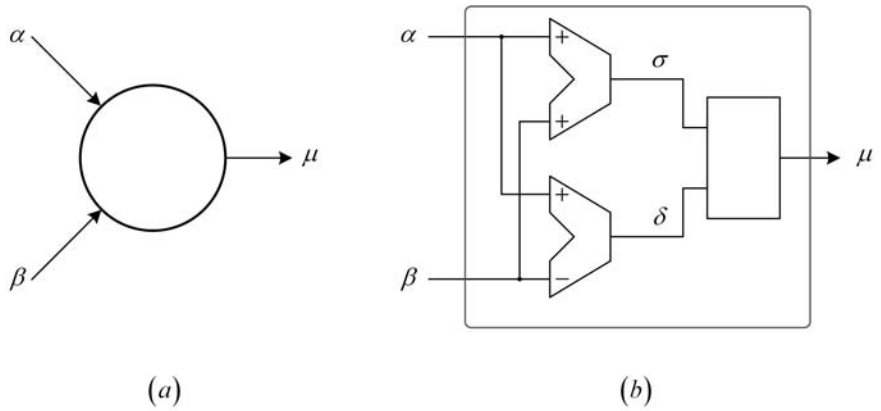

Abbildung 4.2: Symbol (a) und Realisierung (b) der Elementarzelle

Abb. 4.2 zeigt die Elementarzelle als Symbol (a) und mit ihrem inneren Aufbau (b). In Hardware besteht die Elementarzelle, die die Funktion $\mu(\alpha, \beta)$ realisiert, aus einem Addierer für σ und einem Subtrahierer für δ. Addierer und Subtrahierer arbeiten parallel, das Ergebnis μ wird bei geeigneter Maschinendarstellung der Winkel aus σ durch eine einfache von $|\delta|$ gesteuerte Bitmanipulation gebildet. Die Berechnung von μ erfolgt also schnell und einfach. Die Auflösung von μ ist gegenüber α und β um ein Bit verbessert, so dass ein durch Verkettung mehrerer Elementarzellenfunktionen gewonnener Schätzwert $\hat{\varphi}$ gegenüber den Ausgangsdaten ϑ erheblich feiner aufgelöst wird.

Die im vorigen Kapitel besprochenen kartesischen Filter beruhen allesamt auf komplexer Summenbildung für $2N + 1$ aufeinanderfolgende Werte von z^4 bzw. $z^4_{normiert}$. Für diese internen Teiladditionen gelten Assoziativ- und Kommutativgesetz, die beliebige Umordnungen im Sinne einer schnellen und aufwandsminimalen Berechnung erleichtern. Diese Gesetze besagen jedoch nicht, dass in einer formalen Beschreibung nach Gl. (4.1) die Reihenfolge der Argumente $\vartheta(k)$ beliebig geändert werden könnte, da bei den verbesserten Konzepten die Summanden abhängig von ihrer zeitlichen Anordnung gewichtet werden müssen.

Für das Drehwinkelkonzept gilt dagegen wie für das Originalkonzept, dass alle Eingangswerte den gleichen Einfluss auf das Endergebnis haben und deshalb auch in beliebiger Anordnung an die Funktion übergeben werden dürfen (verallgemeinertes Kommutativgesetz). Mathematisch gesehen werden beim Dreh-

winkelkonzept die Argumente als Menge behandelt, nicht als Folge wie beim verteilten Winkelschätzer und dem normiert-gewichteten Phasenschätzer. Dieser Unterschied ist nicht unwesentlich, denn die zeitliche Reihenfolge ist eine wichtige Information für die Phasenschätzung. Schließlich geht es um die Verfolgung einer zufälligen Bewegung (*random walk tracking*), nicht um die Auswertung einer ungeordneten Stichprobe.

Bei dem elementarzellenbasierten Konzept wird die Berechnung wie bei den kartesischen Konzepten intern mit Grundoperationen durchgeführt, für die zwar mit einer kleinen Einschränkung das Kommutativgesetz, aber nicht allgemein das Assoziativgesetz gilt. Die Nichtassoziativität bei Filtern nach dem nun folgenden Konzept ist eine bemerkenswerte Eigenschaft, die es nötig macht, dass man verschiedene Winkelfilter untersucht, die auf den ersten Blick äquivalent erscheinen mögen.

4.4 Winkelschätzer-Baumstruktur

Der wahrscheinlichste Verlauf der physikalischen Phase φ ist derjenige mit der geringsten Änderung zwischen aufeinanderfolgenden Werten, solange man von Intradynempfang mit $\omega_{IF} \approx 0$ ausgehen kann. Schnelle Schwankungen innnerhalb der Zeitreihe der $\vartheta(k)$ sind also auf Rauschen oder Messfehler zurückzuführen und sollen vom Phasenschätzer geglättet werden; dies erfolgt bei den kartesischen Filtern für Realteil und Imaginärteil jeweils einzeln durch ein spezielles digitales Filter.

Man kann aber auch den Verlauf der Phase, wie er sich in den Lagewinkeln ϑ darstellt, direkt glätten. Anders als beim Drehwinkelfilter wird bei dem elementarzellenbasierten Konzept grundsätzlich die Reihenfolge beachtet, denn es geht von der Interpolation zwischen zeitlich benachbarten Werten aus.

Die korrekte Mittelwertbildung aus zwei Lagewinkeln wurde bereits hergeleitet. Um aus einer größeren Anzahl von Lagewinkeln einen einzelnen Mittelwert zu erhalten, ist es wie bei der Addition möglich, in einer Baumstruktur, deren Knoten jeweils Elementarzellen zur Berechnung eines Teilergebnisses darstellen, sukzessive und teilweise parallel (durch Knoten auf gleicher Höhe) zum Endergebnis zu gelangen. Für einen Mittelwert von $2N + 1$ gleichrangigen Werten benötigt man einen Binärbaum mit $\lceil ld(2N + 1) \rceil$ Verarbeitungsebenen und $2N$ Berechnungsknoten.

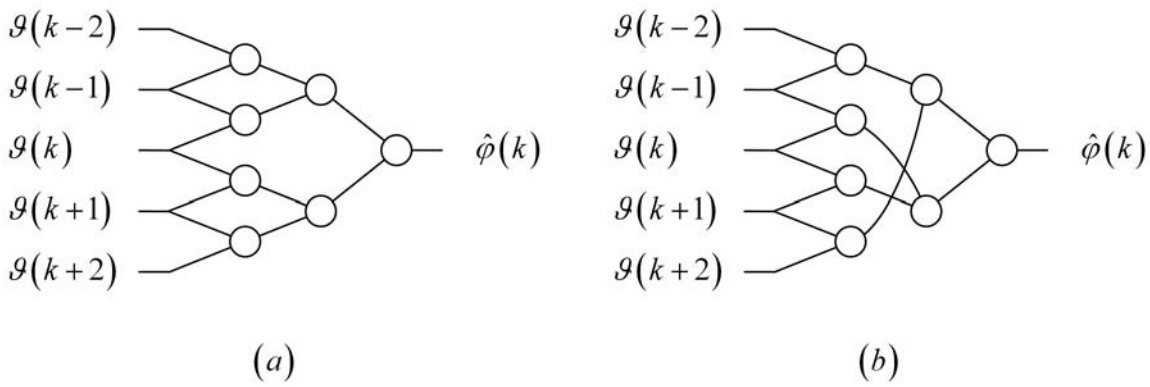

Abbildung 4.3: Zwei Winkelfilter-Baumstrukturen mit N=2

Abb. 4.3 zeigt zwei verschiedene Topologien für Winkelschätzer mit $N = 2$ und sieben Elementarzellen. Eine Elementarzelle liefert aus zwei Winkeln anhand von Summe σ und Differenz δ gemäß (4.6) einen Mittelwert μ, dieser wird mit einem parallel berechneten anderen Wert wiederum gemittelt und so fort bis zur Reduktion auf den einen gesuchten Mittelwert.

Die drei mittleren Werte $\vartheta(k-1), \vartheta(k), \vartheta(k+1)$ werden durch Doppeleinspeisung stärker gewichtet als die Randwerte $\vartheta(k-2)$ und $\vartheta(k+2)$. Dadurch kann zumindest grob die symmetrisch fallende Gewichtung nachgeahmt werden.

Bemerkenswerterweise liefern die Topologien a und b aus Abb. 4.3 im Allgemeinen nicht das gleiche Ergebnis. Dagegen wäre dies bei Addiererbäumen mit den dargestellten Topologien der Fall, denn die Klammersetzung in Summen ist nach dem Assoziativitätsgesetz der Addition beliebig: $(a+b)+(c+d) = (a+d)+(b+c)$. Bei den Elementarzellen wird aber eine komplizierte Operation realisiert, siehe Gl. (4.6) und (4.7), für die kein Assoziativitäsgesetz gilt.

Die Nichtäquivalenz der beiden Topologien lässt sich am einfachsten durch Beispiele nachweisen. Anschaulich klar und praktisch wichtig ist, dass die Mittelung aus zeitlich weit auseinanderliegenden Werten, die in der oberen Elementarzelle in der mittleren Verabeitungsebene (dargestellt als Spalte) in Topologie b stattfindet, für die Phasenschätzung von Nachteil ist. Für die Mittelwertbildung auf dem kürzesten Weg gilt an dieser Stelle als maximal detektierbare Frequenz nur noch ein Drittel der durch den internen Aliaseffekt bestimmten Grenze. Bei den beiden Elementarzellen in der mittleren Spalte von Topologie a gilt dagegegen unverändert f_{max} nach Gl. (3.7).

Die Baumstruktur muss so gewählt werden, dass Mittelungen aus weit auseinanderliegenden Werten erst spät auftreten, wenn die Zwischenwerte schon einen

durch die ersten Stufen geglätteten Verlauf haben. Man kann das Verfahren als sukzessive verteilte Interpolation bezeichnen, bei der eine Zeitreihe von Lagewinkeln über mehrere Zwischenstufen in eine mod$\frac{\pi}{2}$ glatte Zeitreihe von geschätzten Phasenwinkeln überführt wird.

In Hardware werden auf Grund des Demultiplexbetriebes jeweils M Werte von $\hat{\varphi}$ gleichzeitig benötigt, nicht nur ein einziger wie in Abb. 4.3. Die Baumstruktur lässt sich vorteilhaft auf die Bildung von M Ausgangswerten aus $2N + M$ Eingangswerten erweitern, so dass z.B. mit Toplogie a und $M = 8$ lediglich 29 Elmentarzellen benötigt werden, nicht etwa 56. Eine weitere Einsparung ist durch Speicherung von Zwischenwerten aus vorigen Berechnungen möglich, so dass sich der Aufwand auf 24 Elementarzellen reduzieren lässt.

Die Topologien mit $N = 2$ in Abb. 4.3 wurden hier als Beipiele ausgewählt, weil sie übersichtlicher sind als die tatsächlich verwendeten. Die Topologie, deren simulierte BER-Kurve als SMLPA in [HofCOTA] vorgestellt wurde, ist mit $N = 4$ zwar aufwendiger, aber verglichen mit der nur leicht besseren normiert-gewichteten Variante NCF immer noch sehr sparsam. Um Ergebnisse zu erzielen, die dem normiert-gewichtetem Konzept annähernd ebenbürtig sind, musste das verteilte Winkelfilterkonzept (Baumstruktur aus Elementarzellen) allerdings noch durch die Einführung von Zuverlässigkeitsmarken entscheidend verbessert werden.

4.5 Verbesserung durch Zuverlässigkeitsmarken

Die erste Elementarzelle des Winkelfilters bildet wie zuvor beschrieben einen Mittelwert aus zwei Winkeln ϑ bzw. Winkelzahlen w. Falls die Winkeldifferenz allerdings exakt $\pi/4$ beträgt, so ist die Entscheidung für einen Mittelwert nach dem Kriterium des kürzesten Weges nicht eindeutig möglich. Im Falle einer Fehlentscheidung ist der Mittelwert um $\pi/4$ verfälscht, was Fehldekodierungen verursachen kann. Auch wenn die Winkeldifferenz nur annähernd $\pi/4$ beträgt, ist aufgrund von überlagertem Rauschen und Rundungsfehlern mit einem unzuverlässigen Mittelwert zu rechnen. Je spitzer der Differenzwinkel ist, desto zuverlässiger ist dagegen der Mittelwert. Ist der Differenzwinkel deutlich größer als $\pi/4$, so ist der gewonnene Mittelwert wieder zuverlässig, weil in Abb. 4.1 ein Punkt in Bereich 2 vorliegt. Problematisch sind solche Punkte, die nahe an der Grenze zwischen Bereich 1 und 2 liegen.

Weil innerhalb eines Elementarzellen-Winkelschätzers aus Mittelwerten wiederum neue Mittelwerte gebildet werden, um am Ende einen zuverlässigen Schätzwert zu gewinnen, kann es sinnvoll sein, zuverlässige und unzuverlässige Mittelwerte zu unterscheiden. Bei der Weiterverarbeitung sind dann ausschließlich oder zumindest bevorzugt die zuverlässigen Mittelwerte zu verwenden und die unzuverlässigen Zwischenergebnisse ungenutzt zu verwerfen. Dazu muß man bei der Mittelwertbildung eine Zuverlässigkeitsmarke (*reliability bit*, kurz R-Bit) erzeugen und in der folgenden Stufe für eine entsprechende Entscheidung verwerten.

Das letztendlich erfolgreiche Vorgehen zur Bildung einer solchen Zuverlässigkeitsmarke R besteht darin, eine maximale Winkeldifferenz festzulegen, bei der der nach der beschriebenen Methode gebildete neue Winkel gerade noch als verlässlich gelten kann. Die Winkeldifferenz muss zur korrekten Interpolation (d.h. auf dem kürzestem Weg) gemäß (4.6) ohnehin berechnet werden[9].

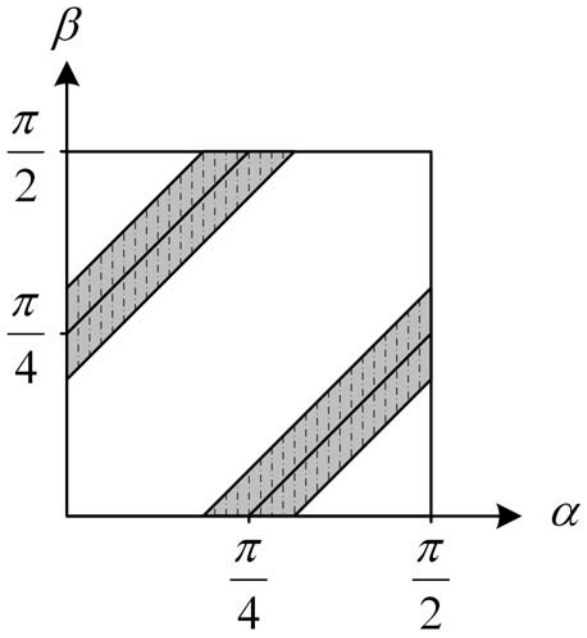

Abbildung 4.4: Bereiche mit $R = 0$

In Abb. 4.4 sind die durch $R = 0$ als unzuverlässig markierten Bereiche durch Schraffur gekennzeichnet. Die Zuverlässigkeitsmarke wird formal durch die folgende Gleichung definiert:

[9]Der Vergleich mit einem solchen Maximalwinkel ist in einigen Fällen (90°, 135°, 180°) besonders einfach realisierbar, weil nur die MSBs der Winkeldifferenz geprüft werden müssen. Ansonsten wäre eine zweite Subtraktion mit anschließender Vorzeichenbetrachtung erforderlich.

$$R = \lceil |\cos\delta| - \cos\delta_R \rceil \tag{4.8}$$

In den Simulationen erwies sich $\delta_R = 3\pi/16$ als das am besten geeignete Kriterium. Die Cosinusfunktionen in Gl. (4.8) dienen lediglich dazu, in der Notation eine kompliziertere Fallunterscheidung zu umgehen. Praktisch wird R direkt aus δ berechnet, δ_R ist eine Konstante.

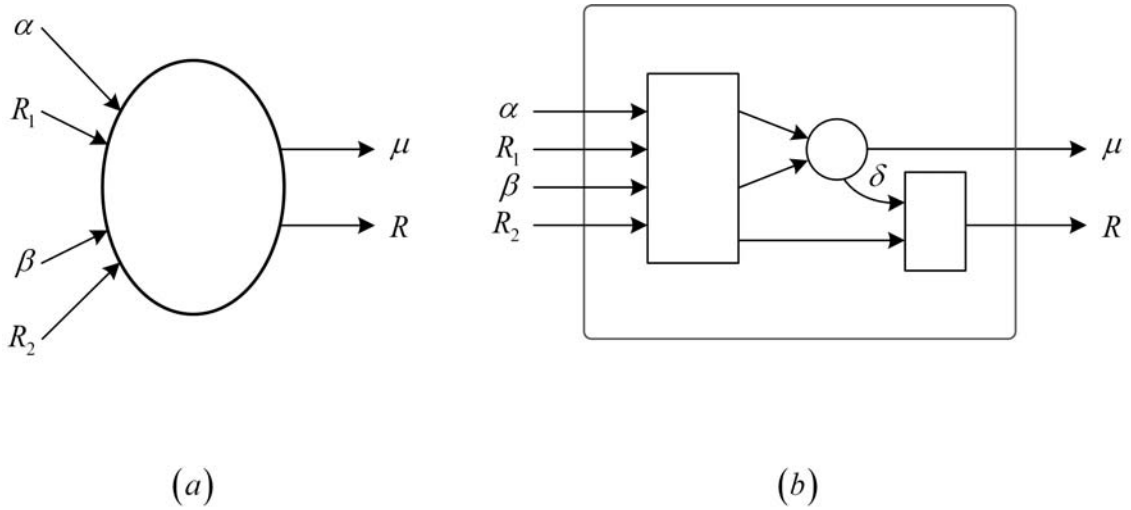

Abbildung 4.5: Symbol (a) und Aufbau (b) der Elementarzelle mit R-Bit

Abb. 4.5 zeigt Symbol und Aufbau einer Elementarzelle, wie sie in der mittleren Verabeitungsebene einer Winkelfilter-Baumstruktur mit Zuverlässigkeitsmarken auftritt. In der ersten Verabeitungsebene gibt es noch keine eingangsseitigen R-Bits, aber die logisch folgende Elementarzelle enthält von den beiden vorgeschalteten Zellen nicht nur deren Interpolationswerte, sondern auch die dazugehörigen R-Bits. Stehen gleich gute ($R_1 = R_2 = 1$) oder gleich schlechte ($R_1 = R_2 = 0$) Mittelwerte zur Verfügung, so wird genau wie in der 1. Stufe ein Mittelwert und ein neues R-Bit für die folgende Stufe gebildet.

Nur bei unterschiedlichen R-Bits kann die Zelle sinnvoll auswählen, indem sie den Eingangswert mit $R = 0$ ignoriert und den jeweils anderen Winkelwert als eigenen Ausgangswert weiterleitet. Praktisch kann letzteres dadurch geschehen, dass man den einen 'guten' Wert mit Multiplexern auf beide Eingänge der eigentlichen Zelle schaltet, denn der Mittelwert von zwei gleichen Zahlen ist mit ihnen identisch. Ein neues R-Bit mit dem Wert 1 wird dabei automatisch erzeugt, weil für zwei gleiche Eingangswinkel $\delta = 0$ ist. Die linke rechteckige Einheit in Abb. 4.5 enthält eine kleine Logikschaltung für die beiden R-Bits und den Multiplexer,

der die Eingänge der eigentlichen Elementarzelle (vgl. Abb. 4.2) ansteuert.

In der Elementarzelle des Winkelfilters werden ohnehin bereits Summe σ und Differenz δ aus den Eingangswinkeln berechnet. Aus δ kann man in der Maschinendarstellung leicht ein Reliabilitätsbit nach der Definition in Gl. (4.8) generieren. Dies ist in Abb. 4.5 (b) durch das kleine Rechteck unten links angedeutet, welches nur eine kleine Logikschaltung enthält.

Bei Festlegung der Verlässlichkeitsgrenzen auf $|\delta| = 3\pi/16$ und $|\delta| = 5\pi/16$ müssen bei geeigneter Maschinendarstellung nur die beiden MSBs des Differenzbetrages geprüft werden. Dazu genügt einfache kombinatorische Logik, ein expliziter Vergleich (also eine zweite Subtraktion) ist nicht erforderlich.

In den folgenden Zellen (Baumknoten) muss jetzt nicht nur mit den Mittelwerten aus den ersten Knoten, sondern auch mit den Zuverlässigkeitsmarken R_1 und R_2 gearbeitet werden, um einen neuen Mittelwert zu erzeugen. Analog zum normierten Vorbild sollte ein Wert μ mit $R_i = 0$ unterdrückt werden, andererseits sollte das Filter auch unter ungünstigen Betriebsbedingungen immer einen Ausgangswert $\hat{\varphi}$ bzw. v für die Dekodierung liefern, um den Rückgriff auf ϑ bzw. w als Ersatzwert (vgl. voriges Kapitel) zu vermeiden. Deshalb wird bei $R_1 = R_2 = 0$ trotzdem ein Mittelwert aus eigentlich schlechten Werten gebildet, der seinerseits sogar wiederum mit $R = 1$ als zuverlässig markiert sein kann.

4.6 Winkelfilter und Gewichtung

Das in Gl. (4.8) vorgestellte Kriterium zur Markierung unzuverlässiger Mittelwerte wurde nicht näher begründet. In der Tat wurden auch andere denkbare Definitionen für R ausprobiert, etwa Berücksichtigung des Betrages der Eingangsdaten z. Wie beim Vergleich von Originalkonzept und normiertem Konzept zeigte sich jedoch, dass $|z|$ keine brauchbaren Informationen zur Zuverlässigkeitsbestimmung liefert.

Zunächst scheint es sich beim Unterscheiden nach R-Bits um etwas ganz anderes zu handeln als die im vorigen Kapitel dargestellte Methode mit Normierung, Gewichtung und Mittelung. Trotzdem gibt es Gemeinsamkeiten, die eine theoretische Erklärung für die gute Funktion des durch Zuverlässigkeitsmarken R erweiterten Winkefilters liefern.

Unterscheiden sich zwei Winkel gerade um π (bzw. bei den QPSK-Lagewinkeln: um $\frac{\pi}{4}$), so kann ihnen nach dem Kriterium des längsten Weges kein eindeutiger

Mittelwert zugeordnet werden, was durch $R = 0$ zu markieren wäre, aber durch eine im Prinzip willkürliche Entscheidung letztendlich gelöst werden muß.

Eine gleich gewichtende Teilsummen-Additionsstufe bei der normierten Lösung verhält sich in diesem kritischen Fall ganz anders: Die Summe aus zwei betragsgleichen komplexen Zahlen, deren Argumente sich exakt um π unterscheiden, beträgt Null und hat also kein definiertes Argument. Allgemein gilt mit $\delta = \alpha - \beta$ für den Betrag der Summe normierter Werte:

$$|e^{j4\alpha} + e^{j4\beta}| = |2\cos 2\delta| \tag{4.9}$$

Das Argument der Summe normierter Werte entspricht 4μ. Allgemeinere Überlegungen zur Addition komplexer Zahlen mit vorgegebenen Beträgen finden sich im Anhang A.3. Die Verwendung des R-Bits kann als Ein-Bit-Ersatzfunktion für die Cosinusfunktion aus Gl. (4.9) aufgefasst werden.

4.7 Ergebnisse zum Winkelfilterkonzept

Der Ansatz, die Tabellierungen, Gewichtungen und komplexe Summenbildung des optimierten Viterbi-Phasenschätzers durch ein Winkelfilter (d.h. eine reelle Ersatzfunktion) zu vermeiden, führte zunächst auf ebenfalls komplizierte Filter mit unbefriedigender Leistung (Drehwinkelfilterkonzept) und auf zwar leichter realisierbare aber auch im Vergleich zum optimierten Viterbi-Phasenschätzer nicht ebenbürtige Filter (Konzept ohne R-Bit). Erst die Einführung des R-Bits führte zu Resultaten, die bei beträchtlich verringertem Aufwand fast ebenso gut arbeiten wie der optimierte Viterbi-Phasenschätzer.

Die entsprechenden Simulationsergebnisse sind in [HofCOTA] dokumentiert: der normiert-gewichtete Ansatz wird dort als *Non-linear Compensation Function* (NCF) bezeichnet, während der direkte Phasenschätzer mit R-Bit als *Selective Maximum Likelihood Phase Approximation* (SMLPA) bezeichnet wird. Mit *Maximum Likelihood* ist dabei die Wahl des kürzesten und zugleich wahrscheinlichsten Weges bei der Mittelung gemeint, während das Adjektiv *Selective* auf die durch Zuverlässigkeitsmarken gesteuerte Selektion unter den Zwischenergebnissen verweist.

Abb. 4.6 zeigt die simulierten Kurven für das beste gefundene normiert-gewichtete Filter (NCF mit $N = 5$, vgl Abb. 3.3) und das beste verteilte Winkelfilter mit R-Bit (SMLPA, $N = 4$). Zum Vergleich sind wiederum die Kurven für

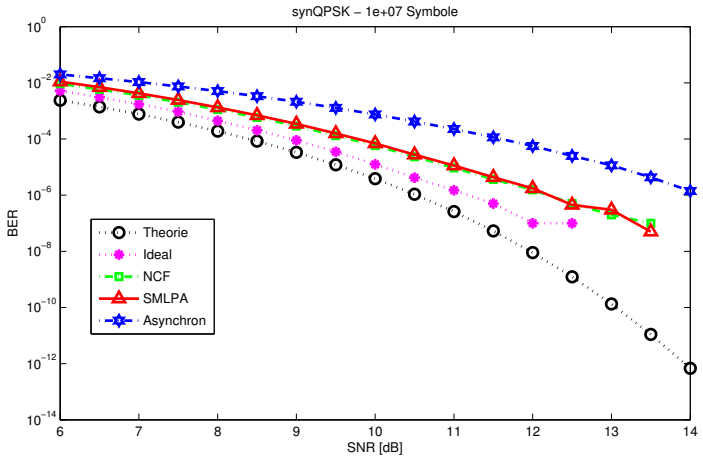

Abbildung 4.6: BER-Kurven mit SMLPA

Asynchronempfang und idealen Empfang (Phasenschätzer, der die tatsächliche Phase kennt) aufgetragen. Ergänzt wurde wie bei [HofCOTA] auch die theoretische Kurve (vgl, z. B. [Proakis]), die Doppelfehler durch differenzielle Kodierung nicht berücksichtigt. Die Übereinstimmung zwischen den BER-Kurven von NCF und SMLPA ist so hoch, dass zur Entscheidung, welches Verfahren besser ist, zusätzliche Berechnungen und Grafiken verwendet wurden (*Ranking*). Dabei zeigte sich, das NCF geringfügig besser ist als SMLPA, aber die Unterschiede waren so gering, dass der erhebliche Mehraufwand zur Realisierung von NCF (vgl. [Romoth07]) nicht gerechtfertigt erschien.

Gleichzeitig mit [HofCOTA] konnten auch schon praktische Resultate mit SMLPA präsentiert werden [PfauCOTA]. Die bei diesem weltweit ersten Echtzeit-Übertragungsexperiment mit DFB-Lasern verwendete VHDL-Umsetzung des Phasenschätzers wurde auch für die CMOS-Implementierung verwendet. Für den Testchip CMOS Version A (Demultiplexer und digitale Signalverarbeitung für einen QPSK-Empfänger ohne Polarisationsmultiplex) wurde eine 120 nm CMOS-Technologie von STMicrosystems (Frankreich) verwendet.

Diese Technologie bietet sechs Leitungsebenen aus Kupfer und verschiedene Bibliotheken mit einfachen und komplexen Logikgattern. Es wurde die HSLL-Variante ausgewählt (*High Speed and Low Leakage*). Der Chip ist pad-limitiert und hat einen Flächenbedarf von 3 mm x 2.4 mm. Der *full-custom*-Teil (Demultiplexer) enthält 6137 Elemente, davon 5145 Transistoren. Die mit 625 MHz getaktete DSPU (*digital signal processing unit*) enthält 63187 Logikgatter. Nach Simulationen besitzt der Chip bei der angestrebten Datenrate von 10 Gbit/s eine Leistungsauf-

nahme von 1.2 W. Der Chip ist produziert worden, Messergebnisse stehen aber noch aus.

Kapitel 5

Polarisationsmultiplex mit digitaler Regelung

5.1 Grundlagen und Modellierung

Im bisherigen Verlauf dieser Arbeit wurde ausgeführt, wie man einen synchronen QPSK-Empfänger digital realisieren kann und insbesondere Lösungsmöglichkeiten für das zentrale Problem der Trägerphasenrückgewinnung vorgestellt. Motivation für den Übergang von der einfachen Intensitätsmodulation zu QPSK war die Verdopplung der Bitrate bei gleichbleibender Symbolrate.

Eine zusätzliche Möglichkeit zur Verdopplung der Symbolrate ist Polarisationsmultiplex: der Senderlaserstrahl wird vor der Modulation in zwei zueinander orthogonale Polarisationen aufgespalten, die jeweils getrennt QPSK-moduliert und anschließend auf einer Faser zusammengeführt werden. Entsprechend müssen auf der Empfängerseite die Polarisationen separiert und einzeln demoduliert werden.

Da beim optischen Überlagerungsempfang auch bei Polarisationsmultiplex nur zwei Laser zum Einsatz kommen, haben die beiden elektrischen Empfangssignale eine gemeinsame IF-Bezugsphase, die mit den in den beiden vorangegangenen Kapiteln beschriebenen Methoden zurückgewonnen bzw. geschätzt werden kann, wobei aber für einen durch den Filterparameter N festgelegten Betrachtungszeitraum die doppelte Datenmenge zur Verfügung steht, nämlich $2(2N+1)$ gegenüber $2N+1$ Werten.

Diese Verdopplung ist leicht in die erfolgreichen Filterkonzepte integrierbar, vgl.

[Romoth07] und bewirkt eine höhere Genauigkeit der Phasenrückgewinnung, die sich als Verbesserung der Bitfehlerraten gegenüber einfachem synQPSK-Betrieb auswirkt.

Die elektronische Synchrondemodulation und Dekodierung erfolgt analog zu der bisher besprochenen. Die Phasen- und Datenrückgewinnungseinheit (PDR, *phase and data recovery*) erhält jeweils pro Modul zwei komplexe Zahlen als Eingangswerte; deren Argumente ψ_1 und ψ_2 werden entsprechend dem Zeitbedarf für die Phasenschätzung verzögert und dann nach Abzug von $\hat{\varphi}$ in empfangene Quadrantenzahlen $n_{r,1}$ und $n_{r,2}$ umgerechnet. Da es nur eine Bezugsphase gibt, gibt es auch nur eine Quadrantensprungzahl n_j, die bei der differenziellen Dekodierung beider paralleler Datenströme verwendet wird.

Die digitalen Eingangsdaten von den vier AD-Umsetzern werden zusammengefasst als zweidimensionaler komplexer oder vierdimensional reeller Vektor $\mathbf{z}$:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \text{ bzw. } \mathbf{z} = \begin{pmatrix} \Re z_1 \\ \Im z_1 \\ \Re z_2 \\ \Im z_2 \end{pmatrix} \tag{5.1}$$

Die Komponenten z_1 und z_2 des Eingangsvektors eignen sich aber nicht direkt für die PDR, denn die beiden senderseitig durch Polarisationsmultiplex getrennten Kanäle erscheinen am elektrischen Empfänger mehr oder weniger miteinander vermischt. Die Übertragung durch die Glasfaser und weitere optische Bauelemente kann durch Multiplikation des senderseitigen Jonesvektors mit einer Jonesmatrix $\mathbf{J}$ beschrieben werden, deren Nebendiagonalelemente i. A. nicht Null sind. So kommt es zum Übersprechen (*crosstalk*) zwischen den beiden Kanälen, was die Fehlerrate des Empfängers erhöht.

Ein einfaches mathematisches Modell für den Eigangsvektor von den AD-Wandlern lautet $\mathbf{z} = \mathbf{J}\mathbf{c}e^{j\varphi}$, in komplexen Komponenten ausgeschrieben

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} e^{j\varphi} \tag{5.2}$$

Dabei sind $c_1$und c_2 unabhängig voneinader gesendete QPSK-Symbole; der skalare Faktor $e^{j\varphi}$ ist der schon bekannte IF-Phasor. Die Koeffizienten der Jonesmatrix $\mathbf{J}$ können gemessen an der Übertragungsrate als langsam verändlich angenommen werden. Diese Modellierung ist geeignet, das Übersprechen zwischen

den beiden Polarisationsmultiplex-Kanälen zu beschreiben. Ebenso lassen sich damit polarisationsabhängige Verluste (PDL, *polarization dependent loss*) beschreiben, nicht jedoch Dispersionseffekte, denn sowohl bei chromatischer Dispersion (CD) als auch bei Polarisationsmodendispersion (PMD) beeinflussen sich die zeitlich aufeinanderfolgenden Symbole gegenseitig, während das verwendete Modell mit einer Jonesmatrix $\mathbf{J}$ zwei gedächtnisfreie Kanäle mit Übersprechen zu beschreiben vermag. Im Rahmen dieser Arbeit ist, wenn von Polarisationsregelung die Rede ist, nur die Kompensation einer Jonesmatrix $\mathbf{J}$ gemeint, keine PMD-Kompensation.

Die formale Lösung des Übersprechproblems besteht darin, dass der empfangene komplexe zweidimensionale Vektor $\mathbf{z}$, bevor er der Phasen- und Datenrückgewinnung zugeführt wird, mit einer komplexen Kompensationsmatrix $\mathbf{M}$ multipliziert wird: $\mathbf{x} = \mathbf{M}\mathbf{z}$. Wenn diese Matrix proportional[1] zur Inversen der Jonesmatrix $\mathbf{J}$ ist, besitzt ihr Produkt $\mathbf{Q} = \mathbf{M}\mathbf{J}$ nur Nebendiagonalelemente, die Null sind, so dass kein Übersprechen mehr stattfindet. Auch polarisationsabhängige Verluste wären in diesem Fall ausgeglichen, weil trotz unterschiedlicher Beträge der Komponenten von $\mathbf{z}$ der kompensierte Vektor $\mathbf{x}$ gleiche Beträge der Komponenten besitzt.

Verschiedene nichtlineare Effekte, z.B. eine nicht auf die idealen 90° abgeglichene Phasenverschiebung im 90°-Hybrid können dazu führen, dass man die Jonesmatrix und entsprechend auch die Kompensationsmatrix nicht als Elemente von $\mathbb{C}^{2x2}$, sondern aus $\mathbb{R}^{4x4}$wählen muss. Das Ergebnis der Kompensation ist in beiden Fällen in Hardware ein Vektor aus vier reellen Größen; es wird aber stets als komplexer Wert interpretiert, weil am Anfang der PDR eine Bestimmung von $\psi_p = \text{arc}(x_p) \text{mod } 2\pi$ erfolgt.

Die Hauptdiagonalelemente von $\mathbf{Q}$ können im Prinzip beliebige komplexe Zahlen ungleich Null sein, insbesondere könnten sie theoretisch auch voneinander verschieden sein, ohne dass die PDR in ihrer Funktion beeinträchtigt wird. Unkompensierte PDL durch unterschiedliche Beträge der Hauptdiagonalelemente von $\mathbf{Q}$ wäre ebensowenig störend wie eine quasi-statische Verdrehung der Komponenten gegeneinander, wie sie aus unterschiedlichen Argumenten resultieren würde. Praktisch bedeutet unkompensierte PDL jedoch, dass im schwächeren Kanal die Quantisierungsfehler bei der Winkelbestimmung erhöht sind, was natürlich unvorteilhaft ist.

[1] proportional bedeutet gleich bis auf einen skalaren Faktor ungleich Null; dieser Faktor darf in diesem Fall auch komplex sein.

Als Sollwert für die Regelung, die $\mathbf{M}$ einstellt, ist daher die einschränkende Festlegung sinnvoll, dass $\mathbf{Q}$ gleich der Einheitsmatrix sein soll[2]. Ein unmittelbar einleuchtender Ansatz zur Verbesserung einer schlecht eingestellten Kompensationsmatrix $\mathbf{M}$ ist die Zuweisung $\mathbf{M_{new}} := \mathbf{Q^{-1}M_{old}}$, die in diesem Zusammenhang als ***Zero-forcing approach*** bezeichnet wird[3]. Da $\mathbf{Q} = \mathbf{M_{old}J}$ ist, würde durch diese Zuweisung direkt $\mathbf{M_{new}J} = \mathbf{Q^{-1}Q} = \mathbf{1}$ erreicht, das neue $\mathbf{Q}$ wäre also die gewünschte Einheitsmatrix.

Dieser Ansatz setzt allerdings voraus, dass man $\mathbf{Q}$ kennt und problemlos invertieren kann. $\mathbf{Q}$ kann aber nicht direkt gemessen werden, nur geschätzt, und eine Matrizeninversion wäre in Hardware sehr aufwendig. Zwar kann man die Inverse einer 2x2-Matrix direkt angeben, es müsste aber dazu im Komplexen die Determinante der Matrix gebildet und jeder Koeffizient durch sie dividiert werden. Die Matrizeninversion lässt sich aber auch näherungsweise durch eine Differenzbildung ersetzen, vgl. [Noe05]. Dadurch muss $\mathbf{M}$ allerdings einerseits mit einer anderen Matrix multipliziert werden, andererseits aber auch additiv (nicht multiplikativ) verändert werden. Die durch einen Abschwächungsfaktor g langsam und stetig verlaufende Änderung von $\mathbf{M}$ erwies sich auch zur Reglereinstellung und Fehlersuche als vorteilhaft.

Die Umsetzung der Polarisationsregelung ist Gegenstand dieses Kapitels. Für die Beschreibung wird die Polarisationsregelung in drei Komponenten aufgeteilt. Grundlage ist ein **Korrelator**, dessen Ergebnis zur **Matrizenaktualisierung** (*update*, dieses Element ist die Polarisationsregelung im engeren Sinne) verwendet wird. Die jeweils aktuelle Matrix wird im zuvor beschriebenen **Kompensator** zur Entzerrung verwendet. Der Kompensator ist also gewissermaßen das Stellglied der Polarisationsregelung. Der Korrelator ist dagegen die Messeinrichtung, mit deren Hilfe erkannt werden muss, ob der Kompensator richtig eingestellt ist und wie diese Einstellung ggf. zu ändern ist.

In Abb. 5.1 sind die genannten Elemente und ihr Zusammenwirken in Form eines Kreisdiagrammes noch einmal vereinfachend dargestellt: der empfangene Vektor $\mathbf{z}$ wird mit einer Kompensationsmatrix $\mathbf{M}$ multipliziert, wodurch der neue Vektor $\mathbf{x} = \mathbf{Mz}$ entsteht. Dieser wiederum dient der Phasen- und Datenrückgewin-

[2]Wobei in Maschinendarstellung die Einheitsmatrix mit einer positiven reellen Konstante multipliziert erscheinen mag, das ist aber nur eine Frage der Festlegung von Stellenwerten bei Integer- oder Fixpunktzahlen.

[3]Gemeinhin wird unter *zeroforcing* verstanden, dass man zur Kompensation einer skalaren Übertragungsfunktion $H(j\omega)$ ihr Ausgangssignal mit ihrem Kehrwert $\frac{1}{H(j\omega)}$ multipliziert; bei Übertragung durch Matrizenmultiplikation wird aus dem Kehrwert die Inverse.

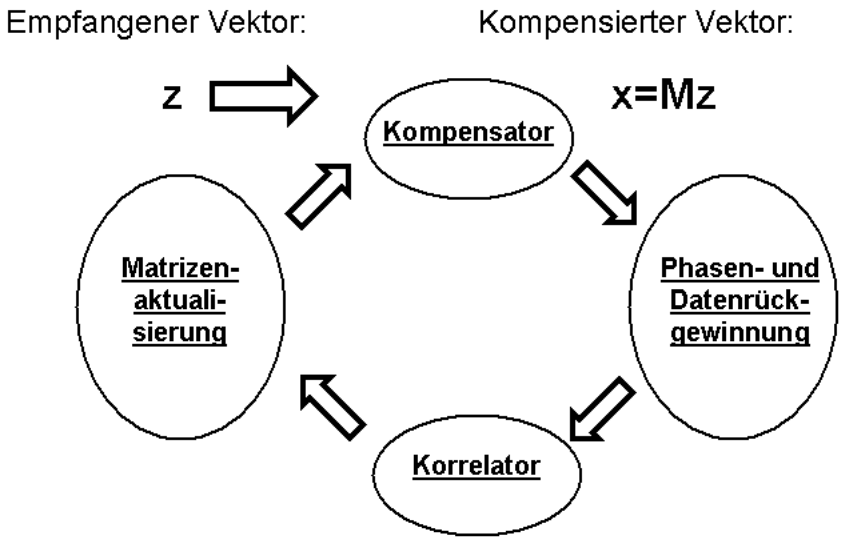

Abbildung 5.1: Elemente der Polarisationsregelung

nung als Eingangswert. Zusammen mit Informationen aus dieser Einheit wird dann im Korrelator festgestellt, ob und in welcher Form es noch zum Übersprechne zwischen den Kanälen kommt. Sind die Komponenten von $\mathbf{x}$ nicht statistisch unabhängig voneinander, so wird bei der nächsten Aktualisierung vom $\mathbf{M}$ entsprechend gegengesteuert. Prinzipiell kann die Polarisationsregelung durch Erreichen von $\mathbf{1} = \mathbf{MJ}$ (Einheitsmatrix) zum Stillstand kommen, durch Rauschen und die Schleifenverzögerung wird die Einheitsmatrix bei dieser Annäherung aber stets knapp verfehlt und die Regelung bleibt in Bewegung.

5.2 Korrelationsbasierte Polarisationsregelung

Im vorigen Abschnitt wurde erläutert, dass nach der Kompensation folgender Vektor vorliegt:

$$\mathbf{x} = \mathbf{Mz} = \mathbf{Qc}e^{j\varphi} \tag{5.3}$$

Optimale Einstellung des Kompensators bedeutet kein Übersprechen, d. h. $q_{12} = q_{21} = 0$. Da die gesendeten QPSK-Symbolfolgen $c_1(k)$, $c_2(k)$ normalerweise un-

korreliert sind, so gilt dies auch für die Komponentenfolgen $x_1(k),\ x_2(k)$. Unkorreliertheit bedeutet für die gesendeten Symbole, dass der Mittelwert ihres Korrelationsproduktes verschwindet:

$$\langle c_1(k) \cdot c_2^*(k) \rangle = \langle c_1^*(k) \cdot c_2(k) \rangle = 0 \tag{5.4}$$

In Gleichung (5.3) erscheint zunächst die gesamte rechte Seite unbekannt, da $\mathbf{Q} = \mathbf{MJ}$ die unbekannte Jonesmatrix enthält und die gesendeten Symbole sowie der IF-Phasor ebenfalls unbekannt sind. Die Phasen- und Datenrückgewinnung liefert aber in Form von $n_{r,1},\ n_{r,2}$ und $\hat{\varphi}$ Schätzwerte, die sich für die korrelationsbasierte Polarisationsregelung benutzen lassen. Multipliziert man beide Seiten von Gleichung (5.3) von rechts mit $\begin{pmatrix} \hat{c}_1^* & \hat{c}_2^* \end{pmatrix} e^{-j\hat{\varphi}}$, so erhält man links ein dyadisches Produkt und rechts ein Produkt von $\mathbf{Q}$ mit dem Skalar $\langle c_1(k) \cdot \hat{c}_1^*(k) \rangle = \langle c_2(k) \cdot \hat{c}_2^*(k) \rangle = 2$ unter der Voraussetzung, dass die Schätzung (Demodulation) immer richtig ist und die Beträge der gesendeten wie der geschätzten Symbole $\sqrt{2}$ betragen, vgl. dazu Gl. (2.6). Das Ergebnis lautet

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} e^{-j\hat{\varphi}} \begin{pmatrix} \hat{c}_1 & \hat{c}_2 \end{pmatrix}^* = \mathbf{Q} \cdot 2 \tag{5.5}$$

Für die linke Seite wird die Bezeichnung $\mathbf{K}$ (Korrelationsmatrix) eingeführt, deren Berechnung im Verlauf dieses Kapitels noch genauer diskutiert wird. Wenn auf die strenge und unrealistsiche Forderung, dass der Entscheider stets das richtige Ergebnis liefert, verzichtet wird, so bleibt die Aussage, dass $\frac{1}{2}\mathbf{K} = \hat{\mathbf{Q}}$ bei hinreichend großer Mittelung ein brauchbarer Schätzer für $\mathbf{Q}$ ist, was die Grundidee für die korrelationsbasierte Polarisationsregelung darstellt. Der Faktor 2 bzw. $\frac{1}{2}$ lässt sich durch entsprechende Skalierung der regenerierten komplexen Symbole vermeiden.

Man müsste den so gewonnenen Schätzwert $\hat{\mathbf{Q}}$ für eine Verbesserung von $\mathbf{M}$ nach dem *zero-forcing-approach* allerdings noch invertieren:

$$\mathbf{M_{new}} := \hat{\mathbf{Q}}^{-1}\mathbf{M_{old}} \tag{5.6}$$

Bei der Matrizenaktualisierung lässt sich die Matrizeninversion jedoch durch eine Approximation vemeiden. Ähnlich der skalaren Approximation $\frac{1}{x} \approx 2 - x$ für $x \approx 1$ gilt für die Inverse einer Matrix $\mathbf{X} \approx \mathbf{1}$die Näherung $\mathbf{X^{-1}} \approx 2 \cdot \mathbf{1} - \mathbf{X}$. Übertragen auf die Aufgabe, eine bereits brauchbare Kompensationsmatrix iterativ zu

verbessern, erhält man

$$\mathbf{M_{new}} := \left(2 \cdot \mathbf{1} - \hat{\mathbf{Q}}\right) \mathbf{M_{old}} = \left(\mathbf{1} - \hat{\mathbf{Q}}\right) \mathbf{M_{old}} + \mathbf{M_{old}} \tag{5.7}$$

Als notwendige Bedingung für die Konvergenz $\mathbf{M_{new}} = \mathbf{M_{old}}$ erhält man insbesondere $k_{12} = k_{21} = 0$, d.h. die Regelung würde erst dann stehen bleiben, wenn die Kreuzkorrelationsprodukte verschwinden. Der Faktor 2 ist wichtig für die numerische Stabilität des Verfahrens. Statt einer direkten multiplikativen Zuweisung von $\mathbf{M_{new}}$ gemäß des mittleren Teils von Gl. (5.7) ist es vorteilhafter, die Kompensationsmatrix in kleinen Schritten additiv zu ändern. Abgeschwächt mit einem skalaren Vorfaktor g wird der in Abhängigkeit von $\hat{\mathbf{Q}}$ veränderte Teil zur alten Kompensationsmatrix addiert:

$$\mathbf{M_{new}} := g \left(\mathbf{1} - \hat{\mathbf{Q}}\right) \mathbf{M_{old}} + \mathbf{M_{old}} \tag{5.8}$$

Das Verfahren wird hier gegenüber [Noe05] und [Noe05JLT] mit leicht veränderter Nomenklatur aber ohne grundsätzliche Änderung dargestellt. In der Hardwareumsetzung ist die in (5.8) doppelt auftauchende Matrix $\mathbf{M_{old}}$ allerdings jeweils unterschiedlich genau maschinell dargestellt, vgl. (5.9).

5.3 Realisierung des Kompensators

Bei den Vorarbeiten zu [Samson06] wurde entschieden, die reelle Variante mit 16 Freiheitsgraden bevorzugt umzusetzen. Auch die wörtliche Umsetzung der komplexen Multiplikation würde 16 reelle Multiplizierer erfordern, nur ist die Kompensationsmatrix $\mathbf{M}$ bei der Variante mit nur 8 Freiheitsgraden stets die reelläquivalente Darstellung einer komplexen Matrix (vgl. A.4). Um den Aufwand für die M-fach parallel erforderliche Matrizenmultiplikation zu begrenzen, verwendet man die Koeffizienten von $\mathbf{M}$ nur mit verringerter Genauigkeit, in der Notation ausgedrückt durch $\mathbf{x} = \mathbf{M_{short}}\mathbf{z}$.

Der Ergebnisvektor $\mathbf{x}$ der Kompensation besitzt auch bei Verwendung von $\mathbf{M_{short}}$ deutlich mehr Binärstellen als der Eingangsdatenvektor $\mathbf{z}$ von den ADCs; deshalb wäre eigentlich für die Argumentbestimmung eine sehr viel größere LUT erforderlich als für den Betrieb ohne Polarisationsmultiplex. Dieses Problem wurde aber durch eine der qmn-Umwandlung nachgeschaltete Bitreduction-Einheit

gelöst [Samson06]. Bei der Bitreduction werden zunächst alle führenden binären Nullen gestrichen, die sowohl in m und n auftreten. Vom Rest der Zahlen m und n wird nur eine feste Anzahl signifikanter Bits verwendet, so dass das gekürzte Zahlenpaar (m, n) eine feste Stellenzahl hat und als Index für eine LUT verwendbar ist.

Nachdem die grundsätzlichen Implementierungsprobleme der Kompensation wie auch der übrigen der Polarisationsreglung in [Samson06] gelöst wurden, sollte in [Wörde07] das System weiter optimiert werden. Bei der Kompensation bedeutete dies, dass für die komplexe Variante ein Kompensator mit verringertem Hardwareaufwand konstruiert werden sollte.

Falls die Kompensationsmatrix komplex ist, gibt es neben der Verwendung der reell-äquivalenten Darstellung, die wie die reelle Lösung 16 Multiplizierer erfordert, die Möglichkeit den Aufwand pro Matrizenprodukt auf 12 Multiplizierer zu reduzieren, vgl. A.5. Die dafür erforderlichen Additionen und Subtraktionen der Matrizenkoeffizienten brauchen nur einmal für alle parallel arbeitenden Module berechnet werden, so dass der Einsparung von $4M$ Multiplizierern nur eine geringer Zusatzaufwand gegenübersteht. Ein weiterer Vorteil ist, dass die für die Multiplikation modifizierte Matrix auch innerhalb der Matrizenaktualisierung verwendet werden kann.

5.4 Matrizenaktualisierung und Optimierung

Die reelle Kompensationsmatrix mit 16 voneinander unabhängigen Koeffizienten wird in der Matrizenaktualisierungseinheit (*update*) für die abschließende Addition mit voller Stellengenauigkeit benötigt, ihre Bezeichnung lautet $\mathbf{M_{long}}$. Die gekürzte Variante $\mathbf{M_{short}}$ wird aber auch intern verwendet, weil eine Multiplikation mit $\langle \mathbf{1} - \mathbf{K} \rangle$ erforderlich ist. Das Ergebnis dieser Operation wird vor der Addition mit $\mathbf{M_{long}}$ mit dem abschwächenden Faktor g gewichtet, so dass sich von einem *update*-Schritt zum nächsten die Kompensationsmatrix nur geringfügig ändert. Da sich bei einer leichten Änderung von $\mathbf{M_{long}}$ nicht jedes Mal auch $\mathbf{M_{short}}$ ändert, besitzt die Polarisationsregelung eine variable Totzeit. An diversen Stellen der Aktualisierungseinheit sind Rundungen, Begrenzungen und Umskalierungen nötig, die dazu führen, dass die ursprünglichen Ergebnisse von [Noe05] nicht exakt übertragen werden können. Als Hilfsmittel für die Modulentwicklung in [Romoth07, Samson06, Wörde07] wurde das Matlab-

Simualtionsprogramm zur Polarisationsregelung in Teilfunktionen zerlegt, die bei einer *bottom-up*-Entwicklung nach und nach in VHDL umgesetzt wurden.

Falls die komplexe Beschreibung des Übertragungssystems ausreichend ist, so ergeben sich mehrere Einsparungsmöglichkeiten für die (dann ebenfalls komplexe) Polarisationsregelung:

Die Kompensation kann mit lediglich 12 reellen Multiplizierern pro Modul durchgeführt werden, indem die jeweiligen komplexen Multiplikationen nach dem Strassenalgorithmus mit 3 Multiplizieren realisiert werden, vgl. Anhang A.5. Die zusätzlich erforderlichen Additionen und Subtraktionen brauchen, insofern sie Koeffizienten von $\mathbf{M_{short}}$ betreffen, nur einmal für alle Module durchgeführt zu werden, und zwar zweckmässigerweise in der Matrizenaktualisierungeinheit.

In der Matrizenaktualisierung können ebenfalls einige Berechnungen vereinfacht werden; bei der komplexen Multiplikation der Kompensationsmatrix $\mathbf{M_{short}}$ mit der gemittelten Korrelationsmatrix können die für die Kompensation gebildeten Summen und Differenzen ebenfalls verwendet werden. Zur Speicherung von $\mathbf{M_{long}}$ genügen 8 reelle Zahlen; bei der i. A. doppelten Verwendung in der Verarbeitung müssen lediglich teilweise die Vorzeichen geändert werden.

Beim Korrelator kann man sich auf jeden Fall auf die Mittelung von 8 Werten beschränken; evtl. kann man auch die Anzahl der Multiplikationen beim dyadischen Produkt von 16 auf 8 reduzieren (1. und 3. Spalte benutzen, 2. und 4. als redundant betrachten), ansonsten müssen jeweils Paare reeller Produkte vor der Mittelung vorzeichenrichtig addiert werden. Beide Varianten wurden im Rahmen von [Wörde07] in VHDL implementiert und im Laborversuch getestet.

Die optimierte komplexe Polarisationsregelung lieferte bei FPGA-Versuchen ähnlich gute Ergebnisse wie die aufwendigere reelle. Trotzdem wurde für die CMOS-Implementation sicherheitshalber die bewährte aufwendigere Variante ausgewählt. Von den Optimerungen aus [Wörde07] ist das serielle Matrixupdate in die CMOS-Implemenatation eingeflossen, die eine wesentliche Flächenersparnis bewirkt hat.

Bisher wurde bei der Matrizenaktualisierung implizit vorausgesetzt, dass die umfangreichen Rechenoperationen (Matrizenaddition und Matrizenmultiplikation) jeweils in einem Schritt komplett erfolgen. Dies entspricht der Realisierung bei [Samson06] (vgl. S.53, Formel 5-30):

$$\mathbf{M_{long,new}} = g\left[\langle \mathbf{1} - \mathbf{K} \rangle \, \mathbf{M_{short,old}}\right] + \mathbf{M_{long,old}} \tag{5.9}$$

Bei den Rechenoperationen dazwischengeschaltete Rundungs-, Begrenzungs- und Skalierungsfunktionen müssen dementsprechend auch parallel arbeiten.

5.5 Realisierung des Korrelators

Die Polarisationsreglung benötigt ein mehrdimensionales Eingangssignal, das in Anlehnung an die Terminologie der Regelungstechnik als Regelabweichung bzw. Regeldifferenz[4] zu bezeichnen wäre. Die Veränderung der Kompensationsmatrix durch Addition eines kleinen aus der Regeldifferenz abgeleiteten Korrekturwertes kann nicht nur als Approximation einer Matrizeninversion interpretiert werden, wie sie der Zeroforcing-Ansatz eigentlich fordert, sondern auch direkt als ein zeit- und wertdiskreter mehrdimensionaler integrierender Regler. Die Koeffizienten der Kompensationsmatrix sind die Zustandsvariablen eines enstsprechenden Systems, dessen Eingangsgrößen vor der Differenzbildung durch ein Korrelationsverfahren gewonnen werden müssen.

Das Korrelationsverfahren besteht im wesentlichen aus der Multiplikation von drei teilweise vektoriellen Größen:

$$\mathbf{K} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} e^{-j\hat{\varphi}} \begin{pmatrix} \hat{c}_1 & \hat{c}_2 \end{pmatrix}^* \tag{5.10}$$

Dieser dyadischen Multiplikation vorgeschaltet sind die LUTs zur Bildung eines Teils dieser Größen aus anderen; nachgeschaltet ist eine Mittelwertbildung (*averaging; integrate and dump*). Die Mittelwertbildung ließe sich im Prinzip auch vermeiden oder reduzieren, weil durch das integrierende Verhalten des Reglers ebenfalls eine Mittelung stattfindet. Diese wurde zur Erleichterung von Fehlersuche und Reglereinstellung implementiert.

Ausgehend von dieser verallgemeinerten Beschreibung des Korrelationsverfahrens werden drei verschiedene Möglichkeiten zur Umsetzung beschrieben:

[4] nach DIN 19226 ist die Regelabweichung gleich der Regelgröße minus der Führungsgröße, während die Regeldifferenz die Führungsgröße minus der Regelgröße ist. Konstante mehrdimensionale Führungsgröße (Sollwert) für die gemittelte Korrelationsmatrix wäre die Einheitsmatrix.

Demodulationsansatz, Remodulationsansatz, potenziell multipliziererfreier Ansatz[5] in Polarkoordinatendarstellung.

Beim Demodulationsansatz [Noe05, Noe05JLT] wird der Vektor $\mathbf{x}$, also das Ergebnis der Kompensation parallel zur winkelbasierten Demodulation im Rahmen der PDR ein zweites Mal demoduliert, und zwar durch einen komplexen Phasor, dessen Argument die wiedergewonnene Phase $\hat{\varphi}$ mit negativem Vorzeichen ist. Da zumindest im Mittel $e^{j(\varphi-\hat{\varphi})} = 1$ gilt, entspricht das Ergebnis dieser Demodulation dem gesendeten Symbolvektor multipliziert mit den Matrizen $\mathbf{J}$ und $\mathbf{M}$.

Die vorrangige Zusammenfassung der ersten beiden Faktoren in dem Ausdruck $\mathbf{x}e^{-j\hat{\varphi}}\hat{\mathbf{c}}^{\mathbf{T}*}$ beim Demodulationsansatz entspricht der theoretischen Begründung, praktisch gesehen kann man aber auch zuerst $\hat{\mathbf{x}}^{\mathbf{T}*} = e^{-j\hat{\varphi}}\hat{\mathbf{c}}^{\mathbf{T}*}$ und dann das dyadische Produkt aus $\mathbf{x}$ und $\hat{\mathbf{x}}^{\mathbf{T}*}$bilden. Diese Zusammenfassung wird als Remodulationsansatz bezeichnet, weil das geschätzte Sendesymbol erneut mit der wiedergewonnenen Phase moduliert wird. Mathematisch sollten die beiden Ansätze äquivalent sein, bei der praktischen Umsetzung gab es allerdings diverse Konsistenzprobleme, die mit der Lagedefinition der QSPK-Symbole und dem Wertebereich von $\hat{\varphi}$ zu tun hatten.

Ein dritter Ansatz basiert auf der Annahme, dass die Hauptinformation an den Korrelationskoeffizienten k_{ij} deren Argumente χ_{ij} seien. Ähnlich wie bei der Demodulation mit $\hat{\varphi}$ kann man unter dieser Prämisse auf komplexe Multiplikationen verzichten und auf Additionen und Subtraktionen von Winkeln zurückgreifen, die teilweise ohnehin durchgeführt werden, wie im folgenden Abschnitt im Detail ausführt wird. In der χ-Matrix werden die Argumente der Korrelationsprodukte zusammengefasst:

$$\begin{pmatrix} \chi_{11} & \chi_{12} \\ \chi_{21} & \chi_{22} \end{pmatrix} = \text{arc}\,\mathbf{K} \tag{5.11}$$

Da die Amplitude von $\hat{\mathbf{x}}^{\mathbf{T}*}$fest vorgegeben ist, kann sich das dyadische Produkt nur durch Schwankungen der entsprechenden Amplituden von $\mathbf{x}$ ändern. Muss man also den Betrag der Korrelationsprodukte zusätzlich bestimmen, so kann man alternativ auch $|x_1|$ und $|x_2|$ bilden.

[5] der winkelbasierte Ansatz selbst ist multipliziererfrei; baut man die Amplitudenbestimmung ein, so benötigt man allerdings wieder Multiplizierer.

5.6 Hardwareeffiziente Berechnung der χ-Matrix

Die Winkel $\psi_p = \text{arc}(x_p) \text{mod} 2\pi$ und die geschätzte Phase $\hat{\varphi}$ werden bereits im Demodulator und Dekodierer verwendet. Ein Teil des Demodulators, den man als Entscheider bezeichnen kann, führt die folgende Operation aus, die die empfangene Quadrantenzahl liefert: $n_{r,p} = \left\lfloor \frac{\psi_p - \hat{\varphi}}{\pi/2} + \frac{1}{2} \right\rfloor \bmod 4$. Die empfangene Quadrantenzahl ergibt zusammen mit der geschätzten Phase das Argument des entsprechenden geschätzten Sendesymbols:

$$\begin{pmatrix} \hat{\psi}_1 \\ \hat{\psi}_2 \end{pmatrix} = \text{arc} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} \bmod 2\pi = \begin{pmatrix} n_{r,1} \\ n_{r,2} \end{pmatrix} \frac{\pi}{2} + \hat{\varphi} \tag{5.12}$$

In Hardware ist nur das Zusammenführen von Quadrantenzahlen (als MSBs) und geschätzter Phase nötig, also noch nicht einmal eine Addition. Diese Überlegung führte zum Remodulationsansatz, bei dem zunächst wiedergewonnene Phase und Symbole zusammengeführt und zum Zugriff auf eine LUT verwendet werden, bevor $\mathbf{K}$ als dyadisches Produkt berechnet wird.

Für die direkte Berechnung des Argumentes eines Korrelationsproduktes wie etwa das von x_1 mit $\hat{x}_2^*$ wäre die Differenz von ψ_1 und $\hat{\psi}_2$ zu bilden. Dies ist bereits deutlich einfacher als komplexe Multiplikation. Die Matrix der Argumente der Korrelationsprodukte lässt sich wie folgt direkt berechnen:

$$\begin{pmatrix} \chi_{11} & \chi_{12} \\ \chi_{21} & \chi_{22} \end{pmatrix} = \text{arc} \begin{pmatrix} x_1\hat{x}_1^* & x_1\hat{x}_2^* \\ x_2\hat{x}_1^* & x_2\hat{x}_2^* \end{pmatrix} = \begin{pmatrix} \psi_1 - \hat{\psi}_1 & \psi_1 - \hat{\psi}_2 \\ \psi_2 - \hat{\psi}_1 & \psi_2 - \hat{\psi}_2 \end{pmatrix} \tag{5.13}$$

Die Differenzbildungen brauchen nicht gesondert durchgeführt zu werden, wodurch sich die Berechnung weiter vereinfacht. Die Elemente der Hauptdiagonale fallen quasi als Nebenprodukte im Entscheider an, und die Elemente der Nebendiagonalen hängen von den Hauptdiagonalelementen dergestalt ab, dass sie leicht aus ihnen zu bestimmen sind. Zunächst werden die Hauptdiagonalelemente χ_{pp} betrachtet mit $p \in \{1, 2\}$:

$$\chi_{pp} = \psi_p - \hat{\psi}_p = \psi_p - \left(n_{r,p} \frac{\pi}{2} + \hat{\varphi} \right) = (q_p - n_{r,p}) \frac{\pi}{2} + \vartheta_p - \hat{\varphi} \tag{5.14}$$

Dabei wurde ψ_p entsprechend der Hardwarelösung in Quadrantenzahl q_p und den Lagewinkel $\vartheta_p$aufgespalten. Die Differenz $\vartheta_p - \hat{\varphi}$ bzw. vollständig notiert $\psi_p - \hat{\varphi} = q_p \frac{\pi}{2} + \vartheta_p - \hat{\varphi}$ wird im Entscheider berechnet; $n_{r,p}$wird vom Entscheider

gerade so bestimmt, dass der Betrag von χ_{pp} minimal wird, so dass diese Gröβe immer aus dem Intervall $[\frac{-\pi}{4}, \frac{\pi}{4}]$ stammt.

Die für die Entscheidung berechnete Teildifferenz $\vartheta_p - \hat{\varphi}$ liegt im Intervall $]\frac{-\pi}{2}, \frac{\pi}{2}[$, weil die beiden Winkel jeweils im Intervall $[0, \frac{\pi}{2}[$ liegen; diese Teildifferenz entspricht gerade χ_{pp}mod $\frac{\pi}{2}$. Interpretiert man die Teildifferenz als vorzeichenbehafteten Winkel in Zweierkomplementdarstellung, so erhält man gerade die gesuchte vorzeichenbehaftete Größe χ_{pp}; es ist sozusagen der Rest, der bei der Entscheidung weggerundet wird. Dieser Sachverhalt soll durch das Zeichen $\cong$ ausgedrückt werden: $\chi_{pp} \cong \vartheta_p - \hat{\varphi}$.

Zur Erklärung, wie auch die Nebendiagonalelemente leicht gewonnen werden können, wird zunächst folgende Beziehung zwischen $\hat{\psi}_1$ und $\hat{\psi}_2$ hergeleitet:

$$\hat{\psi}_1 = n_{r,1}\frac{\pi}{2} + \hat{\varphi} = (n_{r,1} - n_{r,2} + n_{r,2})\frac{\pi}{2} + \hat{\varphi} = (n_{r,1} - n_{r,2})\frac{\pi}{2} + \hat{\psi}_2 \tag{5.15}$$

Mit der neu definierten und leicht berechenbaren Gröβe $n_d = n_{r,1} - n_{r,2}$, die wie die anderen bereits eingeführten Quadrantenzahlen modulo 4 zu lesen ist, erhält man also für die Differenz der Argumente des remodulierten Signals:

$$(\hat{\psi}_1 - \hat{\psi}_2)\text{mod } 2\pi = n_d\frac{\pi}{2} \tag{5.16}$$

Sie unterscheiden sich also bemerkenswerterweise stets um ganzzahlige Vielfache von $\frac{\pi}{2}$, es gilt also, dass die komplexen Komponenten $\hat{z}_1$ und $\hat{z}_2$ des remodulierten Signals, als Vektoren in der zweidimensionalen Ebene betrachtet, zueinander entweder parallel ($n_d = 0$), antiparallel ($n_d = 2$) oder orthogonal ($n_d = 1$ oder $n_d = 3$) sind; weitere Fälle treten nicht auf.

Statt eines formalen Beweises sei nur ergänzend darauf hingewiesen, dass diese Eigenschaft aus der Verwendung einer gemeinsamen geschätzten Bezugsphase $\hat{\varphi}$ für beide Polarisationskanäle resultiert, beide remodulierten Signale werden also mit dem gleichen Lagewinkel konstruiert.

Man kann die gesuchten Nebendiagonalememente der χ-Matrix aus den Hauptdiagonalelementen ableiten, indem man $\hat{\psi}_2$ durch $\hat{\psi}_1$ unter Verwendung von n_d substituiert. Das bedeutet für χ_{12}:

$$\chi_{12} = \psi_1 - \hat{\psi}_2 = \psi_1 - (\hat{\psi}_1 - n_d\frac{\pi}{2}) = \chi_{11} + n_d\frac{\pi}{2} \tag{5.17}$$

Für χ_{21}muss man die Indizes 1und 2 vertauschen und das Vorzeichen von n_d ändern, es ergibt sich

$$\chi_{21} = \chi_{22} - n_d \frac{\pi}{2} \tag{5.18}$$

Anders als die Hauptdiagonalelemente können die Nebendiagonalelemente Winkel in allen vier Quadranten darstellen, die soeben dafür hergeleiteten Gleichungen sind mod 2π oder als Hauptwert zu lesen. Die Addition bzw. Subtraktion der 2-Bit-Quadrantenzahl n_d erfordert in Hardware lediglich die Zusammenführung von MSBs und LSBs, keine echte Addition oder Subtraktion.

Die Zusammenfassung von Formel (5.17) und (5.18) in Matrizenschreibweise ergibt:

$$\begin{pmatrix} \chi_{11} & \chi_{12} \\ \chi_{21} & \chi_{22} \end{pmatrix} = \text{arc} \begin{pmatrix} x_1\hat{x}_1^* & x_1\hat{x}_2^* \\ x_2\hat{x}_1^* & x_2\hat{x}_2^* \end{pmatrix} = \begin{pmatrix} \chi_{11} & \chi_{11} + n_d\frac{\pi}{2} \\ \chi_{22} - n_d\frac{\pi}{2} & \chi_{22} \end{pmatrix} \tag{5.19}$$

Die Argumente der Korrelationsmatrix, also die vier Winkelgrößen $\chi_{11}, \chi_{12}, \chi_{21}, \chi_{22}$ werden im Folgenden zusammenfassend mit χ_{ij} bezeichnet. Sie sind aber genau wie die komplexen dyadischen Produkte nach dem ursprünglichen Polarisationsreglungskonzept erst nach einer Mittelung aussagekräftig im Sinne einer Korrelation und damit für die Aktualisierung der Kompensationsmatrix nutzbar.

Geht man von der wortgetreuen Umsetzung der entsprechenden Formeln aus, so müssten aus den χ_{ij} ähnlich wie beim normierten Konzept anhand von Tabellen komplexe Zahlen $e^{j\chi_{ij}}$ gebildet werden, die dann gleichgewichtet[6] aufzusummieren wären, in diesem Fall sinnvollerweise über M (Anzahl der im Multiplexbetrieb parallel arbeitenden Kanäle). Man kann aber die Überlegungen, die bei der normiert-gewichteten Phasenrückgewinnung zu den Winkelschätzerkonzepten führen, ebenfalls auf die χ_{ij} übertragen, wobei allerdings einige Besonderheiten zu beachten sind.

Für χ_{11}und χ_{22} gilt, wie im vorigen Abschnitt ausgeführt, dass sie stets im Intervall $[\frac{-\pi}{4}, \frac{\pi}{4}]$ liegen, was in der Maschinendarstellung analog zu ϑ und $\hat{\varphi}$ problemlos durch $[0, \frac{\pi}{2}[$ ersetzbar ist. Hier wird aber zunächst weiter von vorzeichenbehafteten Winkeln ausgegangen, so dass der Wert 0 in der Intervallmitte liegt. Wichtig

[6]Eine Gewichtung dieser normierten Werte erscheint nicht sinnvoll, denn für den wahren Korrelationskoeffizienten sind alle Einzelwerte gleich gute bzw. schlechte Schätzwerte. Die evtl. aussagekräftigen Beträge der Produkte $z_i\hat{z_j^*}$stehen im hier behandelten Ansatz nicht zur Verfügung.

ist, dass bei einer guten Einstellung der Kompensation der Betrag dieser Winkelgrößen gegen Null streben sollte, so dass der Realteil der Korrelationsprodukte gegen eins geht (Cosinusfunktion).

Ist der Kompensator dagegen schlecht eingestellt, häufen sich Werte von χ_{11}, χ_{22} mit Beträgen am Intervallrand. "Schlecht eingestellt" bedeutet entweder ein Übersprechen zwischen den Kanälen, dass mit den Nebendiagonalelementen erkannt und ausgeregelt werden sollte, oder aber eine Phasenverschiebung gegenüber der Bezugsphase, die durch entsprechende Einstellung von q_{11}, q_{22} korrigiert werden sollte. Die Mittelung der Winkelgrößen χ_{11}, χ_{22} muss ähnlich wie bei der Phasenrückgewinnung mit einer mod $\frac{\pi}{2}$-Mittelung erfolgen, sonst (d.h. bei normaler Mittelung oder Mittelung mit mod 2π) ergibt sich in beiden beschriebenen Situationen ein Mittelwert nahe Null.

Für die Argumente der Nebendiagonalelemente, also χ_{12}, χ_{21}, kann es dagegen Werte aus allen vier Quadranten geben und das Ergebnis ihrer Mittelung liegt auch in $]-\pi, \pi]$ bzw. $[0, 2\pi[$. Auch hier kann ein Binärbaum aus Elementarzellen mit R-Bit-Logik die komplexe Mittelwertbildung ersetzten, die Mittelung muss nur auf diese andere Periodizität bezogen sein. Praktisch ändert sich aber fast nichts, da χ_{12}, χ_{21} gegenüber den Winkeln χ_{11}, χ_{22} zwei Binärstellen mehr besitzen, die der Differenzwinkelzahl n_d entsprechen.

5.7 Winkelbasierte Korrelation und Eindeutigkeit

Die χ-Matrix ist, wie im vorigen Abschnitt dargelegt wurde, sehr leicht zu gewinnen. Unter dem Aspekt der Hardwareersparnis bei der Optimierung des Systems aus [Samson06] durch [Wörde07] wurde dieser Ansatz daher weiter untersucht, aber letztendlich verworfen. Problematisch ist, dass die χ-Matrix für die Verwendung bei der Matrizenaktualisierung wieder in eine komplexe K-Matrix umgerechnet werden muss. Da der Betrag der Elemente von K für die Ausregelung von PDL unverzichtbar ist, hätte darüberhinaus der Betrag der empfangenen und kompensierten Symbole $\mathbf{x}$ gewonnen und mit einem tabellierten Wert für $e^{j\chi_{ij}}$ multipliziert werden müssen. Dieser Aufwand ist nach [Wörde07] größer als bei den konventionellen Konzepten. Eine winkelbasierte Implementierung ohne Verwendung der Amplitudeninformationen wurde im Labor getestet, die Ergebnisse waren aber erwartungsgemäß nicht überzeugend.

Ein weiteres Problem ist das der Eindeutigkeit und Optimalität der Regelung.

Es wurde bereits angesprochen, dass $\mathbf{Q} = \mathbf{MJ}$ nicht unbedingt auf die Einheitsmatrix abgegelichen werden muss, damit Übersprechen verhindert wird und die PDR möglichst fehlerfrei arbeitet (BER als externes Optimalitätskriterium), aber dass die (in Maschinenzahlen passend skalierte) Einheitsmatrix eine sinnvolle Zielvorgabe für die gemittelte Korrelationsmatrix ist.

Das Problem besteht nun darin, dass $\langle \mathbf{K} \rangle = \mathbf{I}$ kein hinreichendes Kriterium dafür ist, dass die Regelung sich optimal verhält in dem Sinne, dass kein Übersprechen stattfindet ($q_{12} = q_{21} = 0$). Falls aber $\langle \mathbf{K} \rangle = \mathbf{I}$ ist, so führt die Aktualisierung zu $\mathbf{M_{long,new}} = \mathbf{M_{long,old}}$, d. h. die Regelung bleibt auf einem nicht optimalen Wert stehen. Dies wurde mit einem vereinfachten und idealisierten Simulationsmodell demonstriert[7], im Laborexperiment aber bisher nicht beobachtet. Abb. 5.2 zeigt

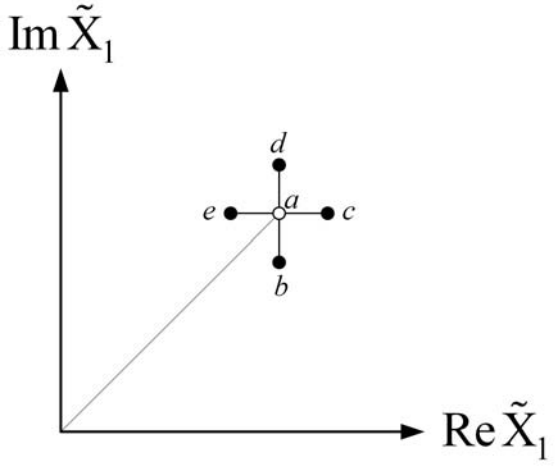

Abbildung 5.2: Suboptimale Einrastung der Polarisationsregelung

verschiedene Werte der Variablen $\tilde{X}_1$ zur Verdeutlichung des Einrastphänomens, die mit den bisher eingeführten Größen x_1, ψ_1 und $\hat{\varphi}$ wie folgt definiert ist:

$$\tilde{X}_1 = |x_1| \exp\left(\left[\psi_1 - \hat{\varphi} + \frac{\pi}{4}\right] \bmod \frac{\pi}{2}\right) \tag{5.20}$$

Verglichen mit dem entsprechenden Korrelationsprodukt K_{11} ist $\tilde{X}_1$ also auf die Standardposition ($\vartheta = \frac{\pi}{4}$) gedreht. Ist der Regler gut eingestellt und das Rauschen gering, so liegen alle Werte von $\tilde{X}_1$ gehäuft um die Position a aus Abb. 5.2. Ist der Regler dagegen schlecht eingestellt, so dass es zum Übersprechen kommt, so hängt die Position von $\tilde{X}_1$ auch davon ab, welches Symbol auf dem zweiten Kanal gesendete wurde.

[7] interne Mitteilung von Prof. Noe

Den vier möglichen QPSK-Symbolen entsprechen durch $q_{12} \neq 0$ vier verschiedene Positionen b, c, d, e von $\tilde{X}_1$. Werden aber alle vier möglichen QPSK-Symbole etwa gleichoft gesendet und ist die Anordnung der Punkte b, c, d, e derart symmetrisch wie in Abb. 5.2, dann ist ein Mittelwert über genügend viele $\tilde{X}_1$ vom Mittelwert bei guter Einstellung nicht zu unterscheiden, beide liegen auf Position a (anschaulich ist der Mittelwert der Schwerpunkt einer Menge von Punkten).

Wie bereits erwähnt konnte ein Einrasten der Polarisationregelung im Laborversuch bisher nicht beobachtet werden. Ursache für das Nichtauftreten der suboptimalen Zustände könnte aber sein, dass die aktuelle FPGA-Implementierung durch die bislang bevorzugte hohe Regelgeschwindigkeit [PPSECOC07, OFC2008] zu Überschwingungen und starkem Eigenrauschen neigt, so dass zwar das Einrastproblem nicht auftritt, aber die Reglung sich nicht optimal verhält. Es ist zu befürchten, dass bei zukünftigen Versuchen mit einer auf Genauigkeit statt auf Geschwindigkeit optimierten Regelung das Einrastproblem doch auftritt.

Für Testchip CMOS B wurden deshalb einige Zusatzfunktionen vorgesehen und implementiert [Wörde07], die das Erkennen und Beheben dieser unerwünschten Zustände erlauben sollen. Das für diesen Chip ohnehin vorgesehene *debug interface* ermöglicht insbesondere die externe Bildung eines Histogramms aus Korrelationsdaten und die Durchbrechung einer suboptimalen Einrastung durch Neuinitalisierung von M mit wechselnden Startwerten.

5.8 Experimentelle Ergebnisse

Die bei Abschluss dieser Arbeit aktuellste Veröffentlichung mit experimentellen Ergebnissen ist [OFC2008]. Die Polarisationsregelung wurde bislang unter dem Gesichtspunkt der maximalen Regelgeschwindigkeit optimiert, wobei bemerkenswerte Erfolge erzielt wurden.

Für die in [OFC2008] dargestellten Experimente wurde ein am Fachgebiet Optische Nachrichtentechnik neu entwickelter elektromechanischer Polarisationsverwürfler (*polarization scrambler*) eingesetzt, der auch für andere Experimente schnelle zeitliche Veränderungen der Jonesmatrix **J** erzeugt. Gemessen wird die Geschwindigkeit dieser Änderungen in rad/s, was sich auf die Darstellung von Polarisationszuständen auf der Poincarékugel bezieht.

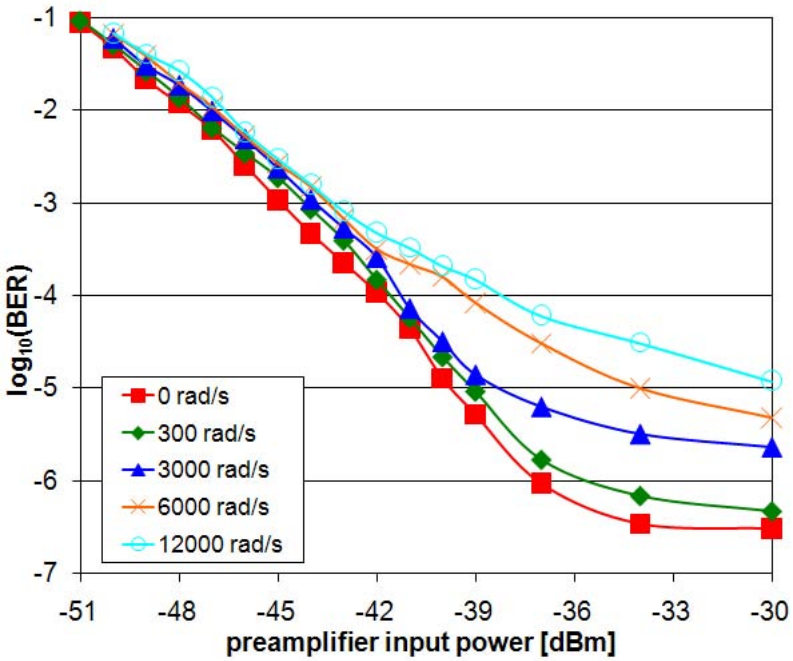

Abbildung 5.3: Schnelle Polarisationsregelung

Abb. 5.3 zeigt die Abhängigkeit der BER von der Eingangsleistung in den optischen Verstärker im Empfänger und damit vom Signal-Rausch-Verhältnis (SNR). Zusätzlicher Parameter der Kurven ist die Änderungsgeschwindigkeit von **J**. Bis 3000 rad/s ist die Polarisationsregelung in der Lage, die zeitvariante Jonesmatrix ohne gravierende BER-Verschlechterung zu kompensieren. Bei weiteren Messungen wurde auch künstlich erzeugte PDL von 3dB problemlos ausgeregelt, und zwar sowohl einzeln als auch kombiniert mit schnellen zeitlichen Veränderungen von **J**.

Kapitel 6

Zusammenfassung und Ausblick

Um den Übergang von traditionellem 10 Gbit/s Direktempfang zu 40 Gbit/s QPSK-Übertragung mit Polarisationsempfang zu bewerkstelligen, wurden die zentralen Probleme der digitalen Signalverarbeitung behandelt und für das Kernproblem, die Trägerphasenrückgewinnung, verschiedene Lösungsansätze präsentiert. Neben einer systematischen und theoretisch gut begründbaren Verbesserung eines bekannten Verfahrens in Kapitel 3 wurde in Kapitel 4 ein nur heuristisch begründetes Konzept entwickelt, das von einer einfachen Grundidee (Winkelmittelung in Elementarzellen) nach einer kleinen Modikfikation (Zuverlässigkeitsmarken) ausgehend vergleichbar gute Simulationsergebnisse erzielte wie das theoretisch abgesicherte aber weitaus aufwendigere zuerst vorgestellte Verfahren. Das heuristische Verfahren wurde in VHDL umgesetzt und sowohl auf einem CMOS-Testchip als auch auf einem FPGA erfolgreich implementiert. Experimentelle Ergebnisse liegen bislang allerdings erst für die FPGA-Implementierung vor.

Die Kombination der Trägerphasenrückgewinnung mit geregeltem Polarisationsmultiplex wurde ebenfalls erfolgreich durchgeführt (Kapitel 5). Es ist zu erwarten, dass im Jahre 2008 weitere erfolgreiche Experimente auf Basis der hier dokumentierten Arbeiten durchgeführt werden können. Ein anspruchsvolles Folgeprojekt für synQPSK ist bei der EU beantragt, bei dem, falls es genehmigt wird, die hier dargestellten Erkenntnisse von großem Nutzen sein dürften.

Anhang A

Anhang

A.1 Abkürzungen und Symbole

ADC *Analog-Digital Converter,* Analog-Digitalumsetzer

CMOS *Complementary Metal Oxide Semiconductor*

FEC *Forward Error Correction,* Vorwärtsfehlerkorrektur

FIR *Finite Impulse Response,* endliche Impulsantwort

FPGA *Field Programmable Gate Array*, Feldprogrammierte Logikgattermatrix

LUT *Look-up Table,* Speichertabelle

PDL *Polarization Dependent Loss,* Polarisationsabhängige Verluste

PMD *Polarization mode dispersion*, Polarisationsmodendispersion

QPSK *Quadrature Phase-Shift Keying,* Quadratur-Phasenumtastung

Matrizen und Vektoren werden durch **Fettdruck gekennzeichnet** , komplexe Größen werden nicht besonders gekennzeichnet. Die komplexe Konjugation wird durch * markiert. Mit ^ werden Schätzwerte und Entscheiderergebnisse gekennzeichnet.

$c(k)$ gesendetes QPSK-Symbol

$\mathbf{c}(k)$ gesendeter QPSK-Symbolvektor bei Polarisationsmultiplex

g	Gewichtung des 'neuen' Teils bei der Matrizenaktualisierung
g_n	Koeffizienten für gewichtete Mittelung
j	imaginäre Einheit $\sqrt{-1}$
$\mathbf{J}$	Jones-Matrix (Modellierung der Übertragungsstrecke)
k	Index für zeitdiskrete Größen
$\mathbf{K}$	Korrelationsmatrix (dyadisches Produkt)
m, n	nichtnegative Hilfskoordinaten
m_j	Mastersprungzahl, aggregierte Sprungzahl
M	Anzahl paralleler Multiplex-Kanäle
$\mathbf{M}$	Kompensationsmatrix (reell oder komplex)
$\mathbf{M_{long}}$, $\mathbf{M_{short}}$	... mit unterschiedlichen Auflösungen dargestellt
$\mathbf{M_{new}}$, $\mathbf{M_{old}}$	... aufeinanderfolgende Werte bei rekursiver Berechnung
n	Allgemein: Quadrantenzahlen im Demodulator
n_j	Sprung(quadranten)zahl
n_o	Originalquadrantenzahl; Output-Quadrantenzahl
n_r	Empfangene Quadrantenzahl (differenziell kodiert)
n_t	Gesendete Quadrantenzahl (differenziell kodiert)
n_x	Unbekannte Quadrantenzahl
N	Konstante für Phasenschätzer-Filterbreite
p	Polarisation (Index bei Polarisationsmultiplex)
q	Winkel-Quadrantenzahl (die MSBs von qw)
qw	zusammengefasste Winkelzahl, die ψ darstellt
$\mathbf{Q}$	Produktmatrix **MJ**
$\hat{\mathbf{Q}}$	Schätzwert für $\mathbf{Q}$ (aus Korrelation)

R	Reliabilitätsbit (Zuverlässigkeitsmarke)
v	Winkelzahl, die $\hat{\varphi}$ darstellt
w	Lagewinkelzahl, die ϑ darstellt
$\mathbf{x}(k)$	Symbolvektor nach Kompensation: $\mathbf{x} = \mathbf{M}\mathbf{z}$
$\hat{\mathbf{x}}(k)$	Wiedergewonnener Symbolvektor (für Korrelation)
$z(k)$	Empfangenes QPSK-Symbol (mit IF-Phase und AWGN)
$\mathbf{z}(k)$	Empfangener Symbolvektor (Polarisationsmultiplex)
α, β	Eingangswinkel Elementarzelle
$\hat{\gamma}$	Geschätzter Vierquadranten-Phasenwinkel
δ	Differenzwinkel (in der Elementarzelle)
ζ	Hilfsgröße für Frequenzschätzer (Phaseninkrement)
ϑ	Lagewinkel des empfangenen Symbols
μ	Mittelwert (Ausgang der Elementarzelle)
ξ	Winkelgröße für Speichertabellen
χ	Argument komplexer Korrelationsprodukte
π	Kreiszahl (3,141...)
σ	Summenwinkel (in der Elementarzelle)
$\varphi(t)$	Zwischenträger-Phasenwinkel (zeitkontinuierlich)
$\varphi(k)$	Zwischenträger-Phasenwinkel (zeitdiskret)
$\hat{\varphi}$	Geschätzer Einquadranten-Phasenwinkel
ψ	Argument des empfangenen Signals
ω	Kreisfrequenz (von Lasern; Indizes S, LO, IF)

A.2 Eine spezielle Darstellung komplexer Zahlen

Es seien x, y reelle Zahlen und q eine beliebige ganze Zahl. Zu einer vorgegebenen komplexen Zahl $z = x + jy$ wird eine zweite komplexe Zahl $\tilde{z}$ definert durch die Zuweisung $\tilde{z} := j^{-q}z$

Es sei $\tilde{z} = m + jn$, die reellen Zahlen m, n sind also Real- und Imaginärteil von $\tilde{z}$. Mit diesen Festlegungen ergeben sich unmittelbar folgende Eigenschaften:

Umkehroperation, Darstellung von z:

$$j^q(m + jn) = j^q\tilde{z} = j^{q-q}z = z \tag{A.1}$$

Die Zahl z ist durch das Zahlentripel (q, m, n) eindeutig darstellbar, die Abbildung von z auf ein solches Zahlentripel ist jedoch auf unterschiedliche Weise möglich, weil zu jedem geordneten Paar (z, q) ein geordnetes Paar (m, n) existiert, das die Gleichung (A.1) erfüllt.

Betragsinvarianz und Identität der vierten Potenz:

Weil allgemein $|j^q| = 1$ gilt, ist $|z| = |\tilde{z}|$. Ferner gilt wegen $j^{4q} = 1$ die folgende Identität, die die Betragsinvarianz mit einschließt:

$$z^4 = \tilde{z}^4 \tag{A.2}$$

Zusammenhänge zwischen den Argumenten:

Es gilt arc$z$ = arc$\tilde{z} + q\frac{\pi}{2}$, weil $j^q = e^{jq\frac{\pi}{2}}$ ist, und aufgrund der vorstehenden Identität arcz^4 = arc$\tilde{z}^4$. Ein weiterer nützlicher Zusammenhang ist arcz^4 = 4arc$z$ = 4arc$\tilde{z}$.

Das Argument (Arkus, arc) einer komplexen Zahl z ist die Menge aller reellen Zahlen φ, für die gilt $z = |z|\, e^{j\varphi}$ (gängige Definition des Argumentes, z.B. [Bron91], S.508; [Frei93], S.7). Die vorstehenden Gleichungen liefern also nicht unbedingt den Hauptwert des Argumentes oder ein anderes bestimmtes Element dieser Menge, konkrete Berechnungsvorschriften sind weiter unten aufgeführt.

Periodizität:

$z = j^q\tilde{z} = j^{q+4}\tilde{z} = j^{q\,mod4}\tilde{z}$, da $j^4 = e^{j2\pi} = 1$ ein neutraler Faktor ist. Es gibt also nur vier im Ergebnis verschiedene Transformationen, die statt durch $q \in \mathbb{Z}$ ebenso durch $q\,mod4 \in \{0,1,2,3\}$beschrieben werden können. Diese vier möglichen Ergebnisse der Funktion $\tilde{z}(z,q) = j^{-q}z$ entsprechen den vier vierten Wurzeln von z^4, die nach dem Satz von Moivre existieren.

Einbettung der ganzen Zahlen:

Sind x, y beide ganzzahlig, so sind es auch m und n, weil durch die Multiplikation mit j^{-q} sich jeweils nur das Vorzeichen und die Zuordnung zu Real- und Imaginärteil ändern kann. Desgleichen behalten Festkommazahlen ihre Stellenzahl und Meßwerte ihren ursprünglichen Fehler.

Spezielle Wahl von q als Quadrantenzahl:

Nun sei q nicht mehr beliebig, sondern aus $\{0,1,2,3\}$ so gewählt[1], dass für $|z| > 0$ gilt: $m > 0$ und $n \geq 0$, man also nichtnegative Hilfskoordinaten erhält. Nur für $z = 0$ wird $m = 0$ zugelassen, so dass das von q unabhängige Transformationsergebnis $\tilde{z} = 0 + j0$ darstellbar bleibt.

Durch die zusätzlichen Forderungen erhält man weitere Eigenschaften der Transformation. Wenn in dieser Arbeit von der (q,m,n)-Transformation die Rede ist, ist der Spezialfall mit nichtnegativen Hilfskoordinaten m, n gemeint. Als Berechnungsvorschrift werden für den nicht eindeutigen Fall $z = 0$ die Zuweisungen $q := 0$ und $\mathrm{arc}(0) := 0$ festgelegt[2].

Existenz und Eindeutigkeit:

für $|z| > 0$ und $q \in \{0,1,2,3\}$ gibt es genau ein Zahlentripel (q,m,n), das die Gleichung (A.1) erfüllt. Dies folgt aus dem Satz von Moivre: von den vier vierten Wurzeln von z^4 liegt genau eine im ersten Quadranten[3]. Die Zahl q erhält damit

[1] Aufgrund der Periodizität könnte q auch aus einer anderen geeigneten Teilmenge von $\mathbb{Z}$ gewählt werden, die getroffenen Wahl vereinfacht die Darstellung, denn es gilt $q = q\,mod4$.

[2] $\arg(0) := 0$ ist z.B. auch in Matlab in der Funktion angle() implementiert.

[3] Zum ersten Quadranten wird hier die positive reelle Achse, aber nicht die positve imaginäre Achse gezählt. Dies folgt aus den Bedingungen für m und n. Der Nullpunkt gehört zu allen Quadanten.

die Bedeutung einer Quadrantenzahl ($q = 0$ bedeutet, dass z im ersten Quadranten liegt etc.).

Hauptwert des Argumentes von $\tilde{z}$:

Da $\tilde{z}$ durch die Festlegung immmer im ersten Quadraten liegt, gilt $\arg \tilde{z} = \arctan \frac{n}{m} \in [0, \frac{\pi}{2}[$.

In dieser wie in der folgenden Gleichung stehen auf beiden Seiten konkrete Werte, keine Mengen. Deshalb eignen sie sich als Berechnungsvorschriften ohne weitere Fallunterscheidungen.

Berechnungsvorschrift für das Argument von z:

$$\text{arc}z \, mod 2\pi = \arctan \frac{n}{m} + \frac{\pi}{2} q \tag{A.3}$$

Das Argument $\text{arc}z \, mod 2\pi$ von z ist nichtnegativ und kann deshalb in Maschinenendarstellung vorzeichenlos sein. Den Hauptwert des Argumentes erhielte man aus der rechten Seite der vorstehenden Gleichung fast immer[4] mit $q \in \{-2, -1, 0, 1\}$. Man kann aber die eigentlich vorzeichenlose Maschinendarstellung des Argumentes genau so gut im Nachhinein als vorzeichenbehafteten Wert in Zweierkomplementdarstellung betrachten, um den Hauptwert zu erhalten.

Addition und Subtraktion:

Liegen zwei komplexe Zahlen z_1, z_2 in der (q, m, n)-Darstellung vor, so lautet ihre Summe:

$$z_1 + z_2 = j^{q_1}(m_1 + jn_1) + j^{q_2}(m_2 + jn_2) = j^{q_1} m_1 + j^{1+q_1} n_1 + j^{q_2} m_2 + j^{1+q_2} n_2 \tag{A.4}$$

Die (q, m, n)-Darstellung biete also für die Addition keine Rechenvorteile; es ist vielmehr zweckmässig, die Summanden vor der Summation wieder in die gewöhnliche kartesische Darstellung zu bringen. Dasselbe gilt für die Subtraktion.

[4] genaugenommen: ein Zahl aus dem Intervall $[-\pi, \pi[$ statt aus $]-\pi, \pi]$ wie beim Hauptwert.

Multiplikation und Division:

$$z_1 z_2 = j^{q_1+q_2}(m_1 m_2 - n_1 n_2 + j(n_1 m_2 + n_2 m_1)) \tag{A.5}$$

Man erhält daraus direkt $q = (q_1 + q_2) mod 4$ und (m, n) für das Produkt, eine vorherige Rückumwandlung ist nicht unbedingt erforderlich. Vielmehr kann es rechentechnisch von Vorteil sein, vier Produkte nichtnegativer Zahlen zu bilden und die nötige Vorzeichenbehandlung durch die sehr leichte Berechnung von q abzudecken. Dass $m = m_1 m_2 - n_1 n_2$ auch negativ werden kann, ist für eine anschließende Rückumwandlung nach der allgemeinen Formel A.1 ohne Bedeutung. Ähnliches gilt für die Division, die wie üblich durch konjugiert komplexe Erweiterung auf komplexe Multplikation und reelle Division zurückgeführt werden kann; hier kann der Imaginärteil n negativ sein.

A.3 Bestimmung des Argumentes einer Summe komplexer Zahlen mit vorgegebenen Beträgen

Bei der Normierung und Gewichtung wird in etwas verallgemeinerter Form die folgende Berechnung durchgeführt:

$$\gamma = \text{arc} \sum_{i=1}^{I} g_i \exp(j\xi_i) \tag{A.6}$$

Dabei sind die Beträge der Summanden als Gewichte g_i fest vorgegeben und damit bekannt, während die Argumente ξ_i die unabhängigen Variablen darstellen. Als Teilaufgabe dieser Berechnung wird folgende Gleichung betrachtet:

$$y = g_1 \exp(j\xi_1) + g_2 \exp(j\xi_2) \tag{A.7}$$

Der Trivialfall $g_1 = g_2 = 0$ sei ausgeschlossen und es wird o.B.d.A. angenommen, dass $g_1 \geq g_2$ ist[5] . Dann kann man wie folgt ausklammern und substituieren mit $g := g_2/g_1$, $1 > g \geq 0$ und $\delta := \xi_2 - \xi_1$:

$$y = g_1 \exp(j\xi_1)(1 + g \exp(j\delta)) \tag{A.8}$$

[5]Wahl der Indizes, die hier nur der Unterscheidung dienen (keine zeitliche Abfolge)

In der folgenden Herleitung wird zunächst nur der letzte Faktor aus (A.8) weiter untersucht, der nur von den gerade definierten neuen Parametern g (relatives Gewicht des betragskleineren Summanden) und δ (Differenzwinkel) abhängt:

$$y = g_1 \exp(j\xi_1)z,\ z = z(g,\delta) = 1 + g\exp(j\delta) \tag{A.9}$$

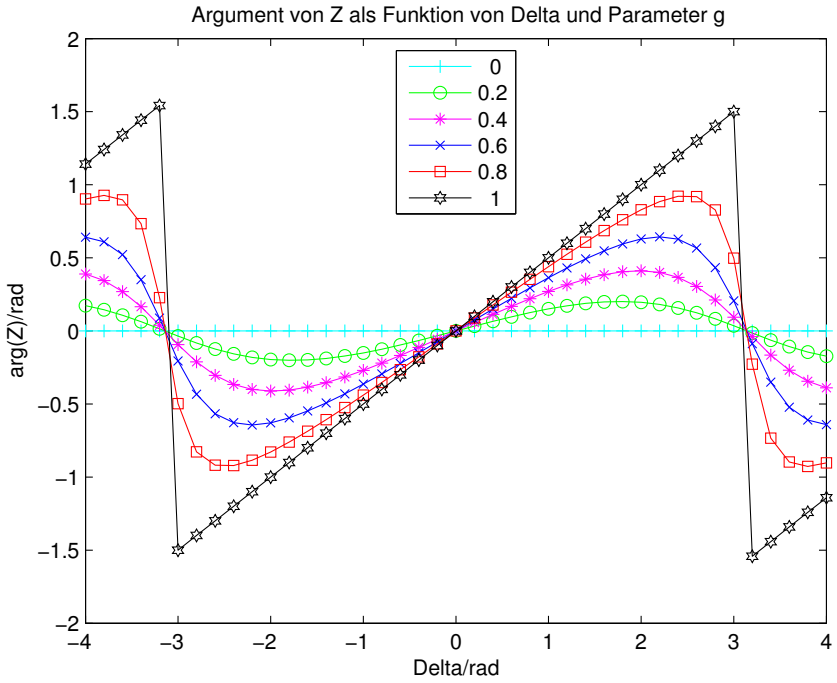

Abbildung A.1: Phasenverlauf von z

Der in Abbildung A.1 dargestellte Phasenverlauf des Faktors z ist für $g < 1$ geschlossen darstellbar:

$$\text{arc}(z) = 2\arctan\frac{\Im z}{|z| + \Re z} = 2\arctan\frac{g\sin\delta}{1 + g\cos\delta + \sqrt{1 + 2g\cos\delta + g^2}} \tag{A.10}$$

Für den Spezialfall $g = 1$ wird diese Funktion stückweise linear, die Unstetigkeitsstellen bei $\cos\delta = -1$ sind in (A.10) beim Grenzübergang $g \to 1$ im Sinne von Dirichlet definiert[6]. Für den $g = 1$ und $|\delta| < \pi$ lässt sich die Phase von z darstellen als

[6]Nämlich als arithmetisches Mittel von rechts- und linksseitigem Grenzwert. Dies entspricht für z der teilweise üblichen aber willkürlichen Festlegung $arg(0) := 0$.

$$\text{arc}(z) = \frac{\delta}{2} \tag{A.11}$$

Dies entspricht der anschaulichen Bedeutung einer Geraden durch den Nullpunkt und z als einer Winkelhalbierenden von δ. Ist $|\delta| > \pi$, so liegt z auf der gegenüberliegenden Seite.

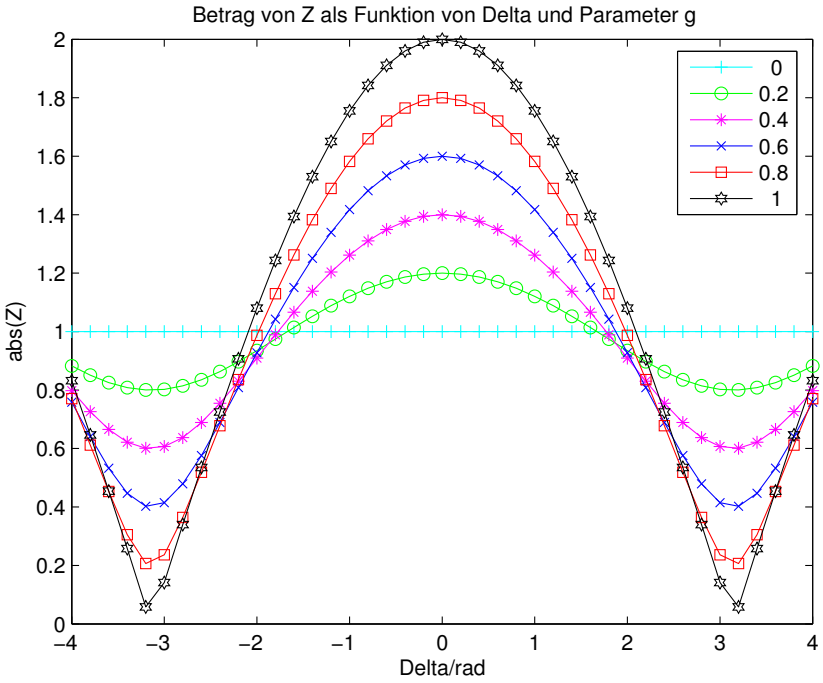

Abbildung A.2: Betrag der Summe: $|Z| = f(\delta, g)$

Der Betrag von z lautet allgemein:

$$|z| = \sqrt{(1 + g\cos\delta)^2 + g^2\sin^2\delta} = \sqrt{1 + 2g\cos\delta + g^2} \tag{A.12}$$

Je größer das relative Gewicht g ist, umso stärker hängt der Betrag von z von der Winkeldiffererenz δ ab. Für kleine Werte von g ist gilt die folgende Näherung:

$$|z| \approx 1 + g\cos\delta \tag{A.13}$$

Für den Grenzfall $g = 1$, also Gleichgewichtung der beiden Eingangswerte, lässt sich (A.12) folgendermaßen vereinfachen:

$$|z| = \sqrt{2 \cdot (1 + \cos\delta)} = 2\,|\cos(\delta/2)| \tag{A.14}$$

Der Fall $g = 1$ ist bedeutsam als Vergleich für die Winkelfilter-Elementarzellen. Die Elementarzellen liefern im Prinzip das Ergebnis in Gl. (A.11), die hier zu Herleitung gemachten Annahmen entfallen dabei.

Um die für z hergeleiteten Formeln für y zu verwenden, muss man gemäß (A.9) resubstituieren und erhält

$$|y| = g_1 \sqrt{1 + 2g\cos\delta + g^2} \tag{A.15}$$

sowie

$$\text{arc}(y) = \xi_1 + \arctan\frac{g\sin\delta}{1 + g\cos\delta} \tag{A.16}$$

A.4 Reelle Beschreibung komplexer Produkte

Die Addition komplexer Zahlen ist bekanntlich äquivalent zur Vektoraddition in $\mathbb{R}^2$, wobei Real- und Imaginärteile der komplexen Summanden als Komponenten zweidimensionaler reeller Vektoren aufgefasst werden:

$$z_1 + z_2 = z \Leftrightarrow (x_1 + x_2) + j(y_1 + y_2) = x + jy \Leftrightarrow \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{A.17}$$

Etwas komplizierter ist es, auch die komplexe Multiplikation in reeller Notation (d. h. ohne Gebrauch der imaginären Einheit j) auszudrücken. Sie kann geometrisch als eine Drehstreckung beschrieben werden, was besonders deutlich wird, wenn man mindestens einen der komplexen Multiplikanden in Polarkoordinaten notiert:

$$z = z_1 z_2 = z_1 \, |z_2| \, (\cos\varphi_2 + j\sin\varphi_2), \varphi_2 = \arg z_2 \tag{A.18}$$

Die Streckung mit $|z_2|$lässt sich in $\mathbb{R}^{\not\varepsilon}$durch einen skalaren Vorfaktor und die Drehung durch eine Drehmatrix ausdrücken. So erhält man folgende äquivalente Notation für $z = z_1 z_2$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = |z_1| \begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \tag{A.19}$$

Auch der zweite komplexe Faktor z_2 kann nach dieser Methode durch eine reelle Matrix ersetzt werden, es sind lediglich z und z_1 als Spaltenvektoren zu notieren. Ersetzt man beide Multiplikanden durch Drehmatrizen mit Betrag, so liegt auch das Ergebnis z als Matrix vor:

$$|z| \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = |z_1| \begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix} |z_2| \begin{pmatrix} \cos\varphi_2 & -\sin\varphi_2 \\ \sin\varphi_2 & \cos\varphi_2 \end{pmatrix} \tag{A.20}$$

Zieht man den skalaren Betragsfaktor in die Matrix und kehrt zu kartesischen Koordinaten zurück, so erhält man folgende Darstellung komplexer Zahlen als reelle 2x2 Matrizen:

$$z = x + jy \Leftrightarrow \mathbf{Z} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \tag{A.21}$$

Diese Darstellung komplexer Zahlen hat gegenüber der Darstellung als zweidiminensionaler Vektor den Vorteil, dass sie mit anderen Operationen von der Addition bis hin zu Matrizenoperationen kompatibel ist und einheitlich gehandhabt werden kann. Mathematisch präziser ausgedrückt: die Menge der Matrizen $\mathbf{Z}$, die Gleichung (A.21) genügen, ist mit der gewöhnlichen reellen Matrizenaddition und -multiplikation als Addtions- und Multplikationsoperation isomorph zu C.

A.5 Berechnung eines komplexen Produktes mit nur drei reellen Multiplikationen nach dem Strassen-Algorithmus

Wenn das Produkt $z \in C$ zweier komplexer Zahlen (gegeben in kartesischer Darstellungsform, $a, b, c, d \in R$) bestimmt werden soll, lässt sich das offensichtlich mit 4 reellen Multiplikationen sowie einer Subtraktion und einer Addition bewerkstelligen:

$$z = (a + jb)(c + jd) = ac - bd + j(bc + ad)$$

Die vier reeellen Produkte ac, bd, bc, ad sind die **essentiellen Terme**. Wenn man versucht, den Aufwand (d. h. vor allem die Anzahl der Multiplizierer) zu verringern, müssen letztendlich die essentiellen Terme auf anderem Wege bestimmt werden. Eine Möglichkeit zur gleichzeitigen Bestimmung mehrer Terme mit nur einem Multiplizierer besteht darin, vor der Multiplikation Summen oder Differenzen zu bilden, z. B.

$$r = a(c + d) = ac + ad$$

Das reelle Produkt r enthält zwei der vier essentiellen Terme. Ein Koeffizientenvergleich mit $\Re z$ liefert

$$\Re z = ac - bd = r - ad - bd = r - d(a + b)$$

Durch eine zweite Hilfsgröße $s = d(a + b)$ kann man $\Re z = r - s$ berechnen. Dadurch ist gegenüber der direkten Berechnung zwar noch nichts gewonnen (Aufwand für den Realteil jeweils zwei Multiplizierer), aber beim Imaginärteil kann man mit einer dritten Hilfsgröße $t = c(b - a)$ ebenfalls r verwenden und so den vierten Multiplizierer vermeiden:

$$\Im z = bc + ad = r + bc - ac = r + t$$

Der Einsparung eines Multiplizierers steht der Aufwand für drei vorgeschaltete Addierer bzw. Subtrahierer gegenüber, welche auch die Berechnungsdauer verlängern. Bedeutung hat das beschriebene Verfahren besonders für große Matrizenmultiplikationen mit starker Besetzung, wo statt einer skalaren Multiplikation eine Multiplikation von Untermatrizen eingespart werden kann, was in der Literatur als Strassen-Algorithmus bekannt ist.

Für eine parallelisierte komplexe Matrizenmultiplikation mit konstanten Koeffizienten (M zweidimensionale komplexe Eingangswerte mal Koeffizienten der Kompensationsmatrix) ist das Verfahren interessant, weil die vorher zu berechnenden Summen mehrfach genutzt werden können, was den zur Einsparung des vierten Multplizierers erfordelichen Mehraufwand insgesamt verringert.

Literaturverzeichnis

[Bron91] I. N. Bronstein, K. A. Semendjajew, Taschenbuch der Mathematik. Moskau/Leipzig 1991 (25. Aufl.)

[Derr] F. Derr, Coherent optical QPSK intradyne system: Concept and digital receiver realization, IEEE JLT, Vol. 10, 1992, pp. 1290-1296

[Fett90] A. Fettweis, Elemente nachrichtentechnischer Systeme. Stuttgart 1990

[Flu06] Fludger C. R. S., Duthel T., Wuth T., Schulien C., "Uncompensated Transmission of 86 Gbit/s Polarization Multiplexd RZ-QPSK over 100 km of NZSF Employing Coherent Equalization", Proc. ECOC'06, Cannes, France, PDP Th4.3.3, pp 33-34.

[Char06] Charlet G. C., Maaref N. M., Renaudier J. R., Mardoyan H. M., Tran P. T., Bigo S. B., "Transmission of 40 Gbit/s QPSK with coherent detection over ultra long haul distance employed by non-linearity mitigation", Proc. ECOC'06, Cannes, France, PDP Th4.3.4, pp 35-36.

[Frei93] E. Freitag, R. Busam, Funktionentheorie. Heidelberg 1993

[Gilb68] B. Gilbert, A precise four-quadrant multiplier with subnanosecond response, IEEE J. Solid-state Circuits, vol. 3, pp. 365-373, 1968

[Gilb82] B. Gilbert, A monolithic microsystem for analog synthesis of trigonometric functions and their inverses, IEEE, J. Solid-state Circuits, vol. SSC-17, pp. 1179-1191, Dec. 1982

[Hess93] W. Hess, Digitale Filter. Stuttgart 1993 (2.Aufl.)

[HofCOTA] S. Hoffmann, T. Pfau, O. Adamczyk, R. Peveling, M. Porrmann, R. Noé, Hardware-Efficient and Phase Noise Tolerant Digital Synchronous QPSK Receiver Concept , CThC6, Coherent OpticalTechnologies and Applications (COTA) Topical Meeting, OSA, Whistler, BC, Canada, June 28-30, 2006.

[HPAECOC07] S. Hoffmann , R. Peveling, O. Adamczyk, T. Pfau, R. Noé: Realtime coherent QPSK transmission: comparison of two carrier phase recovery approaches. ECOC 2007, Berlin, 16.-20. Sept. 2007

[Hooij94] P.W. Hooijmans, Coherent Optical System Design. Chichester 1994

[IpKahn05] E. Ip, J. M. Kahn, Carrier Synchronization for 3- and 4-bit-per-Symbol Optical Transmission, JLT Vol 23, No.12, pp. 4110-4124, Dec. 2005

[IpKahn07] E. Ip, J. M. Kahn, Feedforward carrier recovery for coherent optical communications, JLT, vol. 25, pp. 2675-2692, Sept. 2007

[IpOpEx08] E. Ip, A. P. T. Lau, D. J. F. Barros, J. M. Kahn: Coherent detection in optical fiber systems. OSA Optics Express, Vol. 16. No. 2, 21. Jan 2008

[Jacob94] G. Jacobsen, Noise in Digital Optical transmission Systems. Norwood 1994

[Kamm76] K. D. Kammeyer, Digital Filter Realization in Distributed Arithmetic. Proc. European Conf. on Circuit Theory and Design, Genua 1976

[Kamm96] K. D. Kammeyer, Nachrichtenübertragung. Stuttgart 1996 (2.Aufl.)

[Kay89] S. M. Kay, A Fast and Accurate Single Frequency Estimator, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 37(12), pp. 1987-1989, IEEE, 1989

[Leven06] A. Leven, N Kaneda et al., Real-time implementation of 4.4 Gbit/s QPSK intradyne receiver using field programmable gate array. Electronic Letters, Vol. 42 No. 24

[LevenOFC] A. Leven, N Kaneda et al., Coherent receivers for practical optical communication systems, Proc. OFC, Paper OThK4, Anaheim (2007)

[Ly-Ga05] D.-S. Ly-Gagnon et al., Unrepeated 210-km transmission with coherent detection and digital signal processing of 20-Gbit/s QPSK signal. Proc. OFC, Paper TuL4, Anaheim (2005)

[Ly-Ga06] D.-S. Ly-Gagnon et al., Coherent Detction of Optical Quadrature Phase-Shift Keying Signals with Carrier Phase Estimation. IEEE JLT, vol 24, Jan. 2006, pp.12-21

[Mili05] B. Milivojevic, Study of Optical Differential Phase Shift Keying Transmission Techniques at 40 Gbit/s and beyond. Diss., Paderborn 2005

[Mäusl95] R. Mäusl, Digitale Modulationsverfahren. Heidelberg 1995 (4.Aufl.)

[Noe03] R. Noé, Phase-noise tolerant feedforward carrier recovery concept for baseband-type synchronous QPSK/BPSK receiver, Accepted IASTED WOC, Banff, Canada, July 14-16, 2003

[Noe04] R. Noé, Phase noise tolerant synchronous QPSK receiver concept with digital I&Q baseband processing, Ninth Optoelectronics and Communications Conference/Third International Conference on Optical Internet (OECC/COIN2004), 16C2-5, Yokohama, Japan, July 12-16, 2004

[Noe05] R. Noé, PLL-Free Synchronous QPSK Polarization Multiplex/Diversity Receiver Concept with Digital I&Q Baseband Processing, IEEE Photon. Technol. Lett., Vol. 17, 2005, pp. 887-889

[Noe05JLT] R. Noé, Phase Noise Tolerant Synchronous QPSK/BPSK Baseband-Type Intradyne Receiver Concept with Feedforward Carrier Recovery, IEEE J. Lightwave Technology, Vol. 23, 2005, pp. 802-802

[Noe92] R. Noé, E. Meissner, B. Borchert, H. Rodler: Direct modulation 565 Mb/s PSK experiment with solitary SL-QW-DFB lasers and novel suppresion of the phase transition periods in the carrier

recovery. Proc. ECOC '92, post-deadline paper Th PD 1.5, Vol. 3, pp. 867-870

[NoeCOTA] R. Noe, U. Rückert, Y. Achiam, F. J. Tegude, H. Porte, European "synQPSK" Project: Toward Synchronous Optical Quadrature Phase Shift Keying with DFB Lasers , invited paper CThC4, Coherent Optical Technologies and Applications (COTA) Topical Meeting, OSA, Whistler, BC, Canada, June 28-30, 2006.

[NPAIMS07] R. Noé,T. Pfau, O. Adamczyk, R. Peveling, V. Herath, S. Hoffmann, M. Porrmann, S. K. Ibrahim, S. Bhandare, Real-Time Digital Carrier & Data Recovery for a Synchronous Optical Quadrature Phase Shift Keying Transmission System, Proc. IMS2007, TH2E-01 (invited), June 3-8, 2007, Honolulu, HI, USA.

[OFC2008] T. Pfau, C. Wördehoff, R. Peveling, S. K. Ibrahim, S. Hoffmann, O. Adamczyk, S. Bhandare, M. Porrmann, R. Noé: Ultra-Fast Adaptive Digital Polarization Control in a Realtime Coherent Polarization-Multiplexed QPSK Receiver. OFC 2008, OTuM3

[PfauCOTA] T. Pfau, S. Hoffmann, , R. Peveling, S. Bhandare, S. K. Ibrahim, O. Adamczyk, M. Porrmann, R. Noé, Y. Achiam: Real-time Synchronous QPSK Transmission with Standard DFB Lasers and Digital I&Q Receiver, CThC5, Coherent Optical Technologies and Applications (COTA) Topical Meeting, OSA, Whistler, BC, Canada, June 28-30, 2006.

[PHAPOpEx] T. Pfau, S. Hoffamnn, O. Adamczyk, R. Peveling, V. Herath, M. Porrmann, R. Noé: Coherent optical communications: towards realtime systems at 40 Gbit/s and beyond. OSA Optics express, Vol. 16 No.2, 21. Jan. 2008

[PHPECOC06] T. Pfau, S. Hoffmann, R. Peveling, S. Bhandare, O. Adamczyk, M. Porrmann, R. Noé, Y. Achiam, 1.6 Gbit/s Real-Time Synchronous QPSK Transmission with Standard DFB Lasers, Proc. 32nd European Conference on Optical Communication (ECOC 2006), Cannes, France, 24-28 September 2006.

[PHPEL06] T. Pfau, S. Hoffmann, R. Peveling, S. Ibrahim, O. Adamczyk, M. Porrmann,S. Bhandare, R. Noé, Y. Achiam, Synchronous QPSK

Transmission at 1.6 Gbit/s with Standard DFB Lasers and Real-time Digital Receiver , Electronics Letters, vol. 18, Sept. 2006, pp. 1175-1176

[PHPPTL06] T. Pfau, S. Hoffmann, R. Peveling, S. Bhandare, S. K. Ibrahim, O. Adamczyk, M. Porrmann, R. Noé, Y. Achiam, First Real-Time Data Recovery for Synchronous QPSK Transmission with Standard DFB Lasers, IEEE Photonics Technology Letters, vol. 18, 2006, pp. 1907-1909.

[Plas94] R. van der Plassche, Integrated Analog-to-Digital and Digital-to-Analog Converters. Dordrecht 1994

[PPH07] Pfau, T.; Peveling, R.; Hoffmann, S.; Bhandare, S.; Ibrahim, S.; Sandel, D.; Adamczyk, O.; Porrmann, M.; Noé, R.; Achiam, Y.; Schlieder, D.; Koslovsky, A.; Benarush, Y.; Hauden, Y.; Grossard, N.; Porte, H.:"PDL-Tolerant Real-time Polarization-Multiplexed QPSK Transmission with Digital Coherent Polarization Diversity Receiver", Proc. IMS2007, TH2E-01, June 3-8, 2007, Honolulu, HI, USA

[PPPPTL07] T. Pfau, R. Peveling, H. Porte, Y. Achiam, S. Hoffmann, S. K. Ibrahim, O. Adamczyk, S. Bhandare, D. Sandel, M. Porrmann, R. Noé, "Coherent Digital Polarization Diversity Receiver for Real-Time Polarization-Multiplexed QPSK Transmission at 2.8 Gbit/s", IEEE Photonics TechnologyLetters, Vol. 19, 2007, No. 24, pp. 1988-1990

[PPSECOC07] Pfau, T.; Peveling, R.; Samson, F.; Romoth, J.; Hoffmann, S.; Bhandare, S.; Ibrahim, S.; Sandel, D.; Adamczyk, O.; Porrmann, M.; Noé, R.; Hauden, J.; Grossard, N.; Porte, H.; Schlieder, D.; Koslovsky, A.; Benarush, Y.; Achiam, Y.: Polarization-Multiplexed 2.8 Gbit/s Synchronous QPSK Transmission with Real-Time Digital Polarization Tracking , Proc. ECOC 2007, Berlin, 8.3.3, 16-20 September 2007

[Proakis] J. G. Proakis, Digital communications, Bd.1 New York, 1989 (2. Aufl.)

[Quinn] B. G. Quinn, The Estimation and Tracking of Frequency, Cambridge, 2001

[Romoth07] J. Romoth, Optimierung und Implementierung einer Signalverarbeitungseinheit zur Demodulation von QPSK-Daten. Studienarbeit am FG Schaltungstechnik, Universität Paderborn, 2006

[Ryu95] S. Ryu, Coherent Lightwave Communication Systems. Norwood 1995

[Samson06] F. Samson, Entwicklung einer Polarisationskontrolle für die optische Nachrichtenübertragung. Diplomarbeti am FG Schaltungstechnik (D132), Universität Paderborn, 2006

[Schwarz99] A. Th. Schwarzbacher et al.: Optimisation and implementation of the Arctan Function for the Power Domain, Electronic Circuits and Systems Conference, Bratislava, Slovakia, pp. 33-36, Sept. 99

[Taylor05] M. G. Taylor, Accurate Digital Phase Estimation for Coherent Detection Using a Parallel Digital Processor. ECOC 2005 (Tu 4.2.6) Vol.2, pp.263-264

[Tsuka05] S. Tsukamoto et al.: Coherent Demodulation of 40-Gbit/s Polarization-Multiplexed QPSK Signals with 16-GHz Spacing after 200-km Transmission. OFC , PDP 29

[Viterbi83] A. J. Viterbi, A. M. Viterbi: Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission IEEE Transactions on Information Theory. vol. 29, no. 4, pp. 543- 551, Jul. 1983

[Volder59] J. E. Volder, The CORDIC trigonometric computing technique, IRE Trans. Electron. Comput., vol. EC-8, no. 3, pp. 330-334, Sept. 1959

[WebbHanzo] W. Webb, L. Hanzo: Modern Quadrature Amplitude Modulation. London, 1994

[Weid96] H. Weidenfeller, A. Vlcek, Digitale Modulationsverfahren mit Sinusträger. Heidelberg 1996

[White89] S. A. White, Application of Distributed Arithmetic to Digital Signal Processing: A Tutorial Review. IEEE ASSP Mag., July 1989, p.4-19

[Wiener] N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series. M.I.T. Press, Cambridge, Massachusetts, 1949/1970

[Wörde07] C. Wördehoff, Optimierung einer Polarisationsregelung für die optische Nachrichtenübertragung. Diplomarbeit am FG Schaltungstechnik (D 142), Universität Paderborn 2007